Offering compensation in response to negative online hotel reviews

Master thesis

Marketing

MSc Economics and Business

Yánick Talakua (480070)

480070yt@student.eur.nl

ERASMUS UNIVERSITEIT ROTTERDAM

Supervisor: dr. Ana Scekic

Date: July 12th, 2021

Abstract

Literature has produced mixed results on the exact impact of offering compensation in response to a customer complaint. Additionally, the definite verdict on the offering of compensation in response to negative online hotel reviews is yet to be determined. This study considers the relationship between offering of compensation in a hotel's response to a negative online review, and the booking likelihood and hotel evaluation score of a reader. By construction, it is tested how this relationship depends on the interaction effect between the review's degree of negativity and the reviewer's level of expertise. By experiments done through a survey (between-subjects design), this study tests when and how hotels should offer compensation in response to a negative online review. By formulating four different hypotheses, the outcome of this study is tested for. It is found that offering full compensation will always lead to a higher booking likelihood in comparison to not offering any compensation at all. However, the impact of offering partial compensation, when compared to offering full or no compensation, depends on which review and reviewer is considered. Cash strapped hotels are advised to at least offer partial compensation in response to a very negative review written by a reviewer with a high reputation.

Keywords: negative hotel reviews, compensation, degree of negativity, reviewer expertise, booking likelihood

Contents

List of Tables

List of Figures

1   Introduction

Imagine going back to the 80's and 90's, and asking random people in the street how they have booked their holidays. The most likely answer you are going to get is that they have booked through a travel agency. These booking processes would take quite some time, as it would often take multiple trips to the agency to explore all possible destinations and make a final decision. The amount of information available was fairly limited to for example booklets, knowledge of the salesman and perhaps videos about the destination (Fick and Ritchie, 1991). However, nowadays, due to the rise of the internet, information about a destination is almost unlimited and potentially overwhelming. If wanted, consumers can spend hours and hours researching and dissecting every inch of their potential destination from the comfort of their own homes. New developments as Google Street View, Virtual Reality and YouTube enable consumers to just about visit their destination without actually being there in person. Another important development has been the rise of booking websites, who offer consumers a very extensive catalogue of for example hotels, campsites, restaurants and bars available at the destination of choice (Luca, 2016). It is no longer needed to visit travel agencies, consumers can book their holiday trips online in just a small amount of clicks. These websites often include a wide range of visual and textual content, but one of their most used and prominent features is the user review system which enables customers to voice their opinion about a visited hotel or restaurant.

Increasing influence of online reviews

As consumers continue to buy an increasing number of goods and services from online stores (as booking websites), and lesser from physical offline retail (as travel agencies), they are less exposed to the influence of selling strategies by store personnel (Zhang and Wedel, 2009). In contrast, the influence of online reviews written by customers (or review writers), a form of electronic word of mouth (e-WOM), on purchase decisions of consumers (or readers of the review) has grown accordingly (Chatterjee, 2001; Gruen et al., 2006; Duan et al., 2008; Lee and Youn, 2009). This increasing influence is for example illustrated by the popularity of review systems on booking websites. Online reviews consist of a customer's opinion about a bought product or service written on online platforms as e-commerce websites, social media or independent forums (y Monsuwé et al., 2004). Readers trust and find online reviews by review writers more credible than traditional marketing communication, as for example advertising. In addition, when readers have no first-hand experience with a product or service, they believe that online reviews are a reliable source of information, which can help them to decide if they want to buy the product or service (Akehurst, 2009; Flanagin and Metzger, 2013). Additionally, readers consider such reviews to be more helpful than the product or service information given by a seller (as for example a hotel) and it has been proven that online reviews have a significant effect on product sales (Park and Lee, 2009; Cheung and Thadani, 2012). It is well described by literature that negative/positive online reviews about a product or service tend to lead to lower/higher sales (Chevalier and Mayzlin, 2006; Hu et al., 2008; Ye et al., 2009).

Responding to online hotel reviews

As online reviews have become an important factor in a reader's purchase decision, companies (as the seller) have started to respond to online reviews, as a way to interact with their customers and enhance customer satisfaction and sales (Rose and Blodgett, 2016; Wu et al., 2020). Much established research into the different ways a company can respond to online reviews is primarily focused on one economic sector, namely hospitality and tourism (Ye et al., 2009; Park and Allen, 2013; Xie et al., 2014; Sparks et al., 2016; Phillips et al., 2017; De Pelsmacker et al., 2018). This research is (partially) based on literature which states that online reviews have a strong impact on readers' purchase decisions when considering hospitality and tourism products and services (Xiang and Gretzel, 2010; Hudson and Thal, 2013; Browning et al., 2013). This finding can be explained as it has been shown that 71% of independent travel and hospitality bookings are done online and 95% of travellers read online hotel reviews on booking websites before booking (Schuckert et al., 2015; Ady et al., 2015). While considering the hospitality sector, much literature is focused on responding to negative online reviews (Sparks et al., 2016), as Chan and Guillet (2011) suggest that not responding to negative online reviews will hurt a hotel's future sales. Such result is in line with Hennig-Thurau et al. (2004), who argue that a negative online review will have a longer lasting impact on a company's reputation in comparison to a positive one.

Offering compensation

A call for further research less answered by current literature is whether hotels should offer financial compensation (as for example (partial) refunds or discounts for future stays) to enhance customer satisfaction and sales (Sparks et al., 2016; Li et al., 2017). While most research looks into the immediate effects of psychological compensation offered in a hotel response (as for example an apology or acknowledgement of the customer complaint) results on the impact of financial compensation offered in a response to a negative online hotel review is fairly limited (Sparks and Bradley, 2014). It is well documented that the offering of monetary compensation can have a significant impact on customer complaints in general (Fornell and Wernerfelt, 1987; Hoffman et al., 1995; Mount and Mattila, 2000; Estelami, 2000), but Davidow (2000) argues that offering financial compensation is less important than the way a company handles and intends to solve the customer complaint. Such result is in line with previous findings by Lewis (1983) and Morris (1988), but contrary to findings by Goodwin and Ross (1989) and Conlon and Murray (1996). Such mixed findings illustrate that a definite verdict on the specific impact of compensatory measures is yet to be established.

Perception of a review and its writer

An important element which may impact whether hotels should offer financial compensation to enhance customer evaluations or the booking likelihood of a reader is a reader's perception of an online review. The perception of a reader is formulated as a reader's opinion about a review after reading it. Primary examples that may influence this perception include the characteristics of a review and reviewer (Vermeulen and Seegers, 2009; Sparks et al., 2016; Kwok et al., 2017; De Pelsmacker et al., 2018). To give examples, it is argued that the

characteristics of a review and review writer, as the valence (negative or positive tone) of a review and the credibility (how trusted a reviewer is in the eyes of a consumer) of the reviewer, have a significant impact on how readers perceive online reviews. To be more specific, literature argues that the valence of a review can for example show how helpful a review is in helping to make a purchase decision. For example, a very positively written review is not always considered to be as helpful by readers (Sparks and Browning, 2011; De Pelsmacker et al., 2018). Additionally, the level of expertise of the reviewer often shows how trusted a reviewer is considered to be, as a reviewer with a high reputation is considered to be more trustful in comparison to other reviewers (Sparks et al., 2016; Kwok et al., 2017). Thereby, it is also argued that these characteristics influence how companies should react to online reviews, but also interact among each other. To give an example, it is argued that when readers read a very positive hotel review written by a reviewer with a high reputation (as for example a professional hotel critic), they are more likely to book a stay with the hotel and there is no absolute necessity for hotels to respond (Vermeulen and Seegers, 2009). Contrarily, when readers read a very negative hotel review by a professional hotel critic, they are less likely to book a stay and hotels should quickly respond to try to limit the review's negative impact on sales (Sparks et al., 2016). Thereby, it is argued by Sparks et al. (2016) that, while apologizing in response to a negative review limits the reviews impact on sales, it should be tested if offering compensation in response can limit (or even turn around) the negative impact of the review on hotel sales.

Gap in literature

Three parties are important when considering online hotel reviews, namely the review writers, the responding hotels and the readers of the reviews and responses. These three parties account for this study's independent variable (compensation) and moderator (interaction between characteristics of a review and reviewer). Research into review writers is well covered, as it well known why review writers write reviews and how these writers can be characterized into different categories (Yoo, 2008). It is also extensively shown how characteristics of a review and reviewer influence the perception of a review and it is well known how these characteristics impact the reader's booking likelihood with a hotel. However, although it is widely shown how to textually respond to a negative review, the definite verdict on the offering of compensation in response to negative online hotel reviews is yet to be established (Sparks et al., 2016). Thereby, it is suggested that the offering of compensation could be moderated by the characteristics of an online review and reviewer (Sparks and Bradley 2014, Sparks et al., 2016). Moreover, Sparks et al. (2016) suggest to design a study which looks into interaction effects between reviews, review writers, hotels and readers, where the element of offering compensation should account for how hotels should respond.

Research question

The most important element in this study is the offering of compensation. To account for and quantify the effects of offering compensation on a reader, the dependent variables booking likelihood and hotel evaluation score of a reader are considered. Thereby, this study considers the interaction effect between different types of reviews

(based on the valence) and review writers (based on the credibility) as a moderator. Such reasoning is made, because it is suggested that the relationship between offering compensation and booking likelihood and hotel evaluation of a reader is potentially moderated by this interaction effect. Moreover, the central question to this research is: when and how should hotels offer compensation in response to a negative online review?

To summarize, this study considers hotel reviews, review writers, responding hotels and readers to construct a conceptual framework. *Table 1* shows an overview of all considered elements within this thesis:

Party	Reviews	Writers	Hotels	Readers
Action	Shared opinion about a hotel	Have visited a hotel and write a review	Respond to reviews	Read reviews and responses
Constructed as	Very or less negative written review	Reviewers with a high or low reputation	Offer no, partial or full compensation	Participants in experimental design*

Table 1: An overview of all considered elements within this thesis. (*) Readers of the reviews and responses are simulated for by participants of the survey's within this study's experimental design.

Academic and managerial relevance

As the participants in this study are asked to read reviews and hotel responses, and thereafter are asked to indicate their willingness to book with a hotel, this study is able give a unique perspective on the relation between reviews, review writers, responding hotels and readers. Thereby, this study covers a knowledge gap in literature and provides the hospitality sector with additional understanding of the purchase decisions of customers based on the offering of compensation. Academic wise, results on the offering of compensation in general have produced mixed results and this topic is less examined for in the hospitality sector. Additionally, literature, which considers the effect of offering of compensation on the booking likelihood of a reader by the interaction effect of the degree of negativity of a review and the reviewer expertise, has not yet been conducted by the best knowledge of the writer.

Managerial wise, a better understanding of whether to offer compensation is of key insight when companies allocate (financial) resources to, for example, a customer care or marketing department, as companies should allocate resources as efficiently as possible (Farrell, 1957) and inefficient responding may lead to an unnecessary overspending on budget (Duan et al., 2008). In a perfect world, companies would have enough financial resources to allocate time and budget to each unsatisfied customer. However, companies are cash strapped and have financial targets and should make decisions in when to offer compensation in response or not. Additionally, managers should gain insight in how customers value the offering of compensation, as it is also argued that customers do not always value compensation offered in return (Piehler et al., 2019).

2   Theoretical background

This study's conceptual framework is primarily based on the relationship between the independent variable of offering compensation and the dependent variables booking likelihood and hotel evaluation score. Additionally, the interaction effect between two independent variables (based on different types of reviews and review writers) is considered as moderator. The following sections will discuss how these variables are constructed.

2.1   Independent variables

2.1.1   Offering of compensation

More recent developments into the effect of financial compensation include Liu et al. (2019), who argue that offering compensation has a significant impact on a customer evaluation if a customer complaint is not considered to be severe. Such finding is in line with Piehler et al. (2019), who show that the offering of compensation in combination with an explanation is the most effective way of responding. Companies should judge the intensity of a customer complaint and offer an explanation and compensation accordingly. Both studies call for future research to determine the exact level of compensation that should be offered. In results by Liu et al. (2019), monetary compensation is limited to coupons only, while discount and (partial) refunds are not considered, as suggested by Sparks et al. (2016) and Li et al. (2017). Moreover, as Sparks et al. (2016) and Li et al. (2017) call for further research into the effect of partial (defined as 50% percent refund) and full refunds, this study extends on research by Liu et al. (2019) by considering both. Additionally, as it is argued that offering compensation is not always beneficial to offer due to costs (Piehler et al., 2019), offering no compensation at all is also considered. Hence, in this study offering compensation is constructed as an independent variable which consists of offering none, partial or full compensation in response to a negative online hotel review.

2.1.2   Reviewer expertise

Credibility is a characteristic of a negative review that influences how reviews are perceived by readers (Kwok et al., 2017). It is well described by literature that, whether a review is seen as credible, has a significant impact on whether the review is taken seriously by a reader (Xie et al., 2011; Cheung et al., 2012; Kusumasondjaja et al., 2012). In addition, reviewer expertise is often used as an indication of the credibility of a review, because it has been shown that a reviewer's expertise level impacts the perceived credibility of reviews (Sparks and Browning, 2011; Flanagin and Metzger, 2013; Kwok and Xie, 2016). Extending on such findings, it has also been suggested that reviews written by anonymous (non-verified) writers are less regarded of in comparison to reviews written by verified writers (Jensen et al., 2013).

More recent publications include (Lo and Yao, 2019), who argue that online hotel reviews written by higher rated profiles (so called experts) have a significantly higher level of perceived credibility than reviews written by lower rated profiles (so called amateurs). It should be noted that such outcome is in line with results obtained by Metzger et al. (2010), who state that a positive relationship exists between reputation and endorsement.

Additionally, Lo and Yao (2019) also suggest that negative reviews are found to be more credible than positive reviews, which is in line with results from neuroscience described in Section 2.1.2. Furthermore, review writers would tend to leave a positive review as part of social norms, which would undermine the credibility of positive reviews (Feldman, 1966; Mizerski, 1982; Lo and Yao, 2019).

Results obtained by Lo and Yao (2019) are taken as base for the construction of the independent variable credibility. Therefore, it has been chosen to consider review expertise as independent variable for credibility in this study. Like Lo and Yao (2019), the categorical independent variable of reputation expertise will consist of two levels, namely review writers with a high and low reputation.

Distinction between reviewers with a high or low reputation

To determine if a reviewer has a high or low reputation, most literature base their distinction on the reviewer ranking system present on most booking websites. Where the review ranking system enables reviewers to score hotels, the reviewer ranking systems reflects a user's individual score. Booking websites as TripAdvisor, Booking.com or Yelp use these ranking systems to stimulate content writing. For example, reviewers can earn more points by writing reviews, commenting on reviews and interacting with other reviewers. Thereby, booking websites try to create a thriving online community, who contribute to the total amount of content on the website and make visiting booking websites more interesting for consumers. A more detailed example of the workings of such system can be given when considering a user who has just registered as a member of a booking website. This reviewer is often referred to as either an 'Travel amateur', 'Rookie traveller' or other naming which indicates that the reviewer does not have many points awarded to his or her profile. After posting reviews, commenting and interacting on the review platform, the reviewer will be rewarded points and will rise up the virtual ranks. Reviewers with a high number of points are often referred to as 'Travel expert' or other naming which indicates that the reviewer has a very high number of awarded points attached to his or her profile. Thereby, most booking websites show how many reviews a reviewer has written and how many helpful votes they have received (Luca, 2016).

Jensen et al. (2013) argue that an anonymous reviewer is less regarded of in comparison to a verified user, while Lee and Shin (2014) show that users with a profile picture of a human are more trusted than users without a profile picture at all. All these findings can be explained as when people consider information of strangers (as online reviews), they tend to look for indications to decide if the given information is trustable or not. Thereby, it has been shown that a reviewer ranking system is often used as an indication of trust and helpfulness (Zhang et al., 2010; Chua and Banerjee, 2015).

This study constructs reviewers with a high reputation as having a verified profile containing a profile picture of a human, are described as 'Travel expert' and have a high number of written reviews and received helpful votes. Contrarily, reviewers with a low reputation are specified as anonymous profiles with no profile picture at all, are described as 'Travel amateur' and have a low number of written reviews and received helpful votes. Thereby,

this study assumes that these characteristics are sufficient to represent a noticeable distinction between reviewers with a low or high reputation.

2.1.3 Degree of negativity

One of the most important elements which influences the perception of a review is the valence or tone of the review (Sparks and Browning, 2011; Kwok et al., 2017). As the valence of a review often shows whether a review writer is happy (positive) or unhappy (negative) about his or her purchase, a study which incorporates the perception of online reviews should include a variable which measures the valence of a review (Sparks et al., 2016). Moreover, publications about the effect of valence in reviews, which include Fornell and Wernerfelt (1987), Tax et al. (1998) and Zeelenberg and Pieters (2004), have shown that there is a significant relationship between the valence of a customer's review and customer satisfaction. To illustrate an example, a negative review is often written by a dissatisfied customer.

However, readers react differently depending on the valence of a review. Sparks and Browning (2011) show that negative online hotel reviews have more impact on readers than positive ones. Such found result is in line with Henning-Thurau et al. (2004), who argue that a negative online review will have a longer lasting impact on a company's reputation in comparison to positive one. Therefore, it is suggested that responding to negative online reviews is more beneficial for companies (Hu et al., 2008; Chan and Guillet, 2011; Sparks and Browning, 2011; Sparks et al., 2016). Similar results are found by research in neuroscience, where it has been shown that negative stimuli have a greater impact on humans than positive stimuli (Cacioppo et al., 1986; Anderson et al., 2003). Additionally, it is suggested that the intensity of valence determines how much impact a stimulus has, as for example a very negative stimulus has a greater impact than a less negative stimulus (Cacioppo et al., 1986). Furthermore, Schoenmüller et al. (2019) show that the distribution of online reviews is heavily skewed towards positive reviews, with fewer neutral and negative reviews. It is argued that a significant portion of positive reviews are written by one-time writers, who are asked to share their opinion directly after purchase. Thereby, it is suggested that positive online reviews may be biased due to self-selection.

Therefore, as it has been shown that responding to negative online reviews is more beneficial, the intensity of valence matters and positive reviews may be affected by self-selection, it has been chosen to represent valence of a review as degree of negativity. The categorical independent variable will consists of two levels, namely very negative reviews and less negative reviews.

Distinction between very and less negative online hotel reviews

To label a negative online review as 'very negative' or 'less negative', two elements are considered. First, it has been shown that the overall rating score given by the reviewer is seen as a good indication of the valence of a review. Review systems often include a textual part, where reviewers can write about their experience, and a fixed rating component, which is often presented by a certain amount of awarded stars when considering booking websites. Reviewers can use these stars to 'score' a hotel, where more stars are awarded to a better

performing hotel. Therefore, the awarding of a low amount of stars often coincides with the valence of a review. To give an example, hotels that underperform are more likely to receive a lower amount of stars by unhappy (more negative) visitors, while well performing hotels are more likely to receive a higher amount of stars by happy (more positive) visitors (Mariani, 2018). Additionally, it has been shown by Vermeulen and Seegers (2009) that 1-star reviews often do not include any positive notes about a product or service, while 2-star reviews do often include a (slight) positive one. Thereby, this study assumes that a very negative review coincides with a lower amount of given stars in comparison to a less negatively written review, as readers of the latter are assumed to be more satisfied about the hotel's performance. To be specific, when considering a 5-point rating scale of possible awarded stars, a very negative written review corresponds to one awarded star, while a less negatively written review corresponds to two given stars. It should be noted that the centre of such scale (3 in this example) often denotes the 'neutral' barrier between negative and positive associations.

Second, another element which determines if a negative review is seen as very or less negative is the used language within a review. If a negative review is written in figurative (metaphorical) language, readers exhibit a lower booking likelihood towards a hotel in comparison to a negative review written in literal language. Such finding can be explained as an online review is considered as information given by strangers. In these kinds of social interactions, it is expected to use a more formal literal language, while the use of the more informal figurative language is considered to be inappropriate. However, it is also argued that the use of such language in online reviews has a harmful effect on the reader's perception of the review and the reviewer (Sparks and Bradley, 2014). To give an example, when a negative review containing such language is written by a reviewer with a low reputation, readers are more likely to ignore the review as they do not trust the review and reviewer, and their booking likelihood is less negatively affected (Yin et al, 2020). In comparison, when the same negative review is written by a reviewer with a high reputation, readers are more likely to be convinced of the authenticity of a negative review and their booking likelihood is more negatively affected (Sen and Lerman, 2007; Vermeulen and Seegers, 2009; Sparks and Browning, 2011). Furthermore, it has been shown by Konrod and Danziger (2013) and Liu (2020) that readers consider reviews written by reviewers, who are higher ranked or have a greater reviewer expertise, to be more trustable and helpful than reviews of reviewers who are lower ranked. Additionally, it is argued that very negative and positive reviews include more figurative language than more moderate reviews, from which this study assumes that the language within very negative reviews coincides with figurative speech, while a less negative reviews should correspond to its counterpart, literal speech (Konrod and Danziger, 2013; Liu, 2020).

Another important element within negative reviews is the use of strong emotional language instead of rational language. To give an example, readers find a review which neatly complains about hotel performance more helpful in comparison to a review which furiously complains (Craciun and Moore, 2019). Yin et al. (2020) explains such finding as readers see hot-blooded or angered language as signs of inadequate reasoning and irrationality, but also as a sign of a very upset customer. Hence, this study assumes that the use of strong emotional language coincides with a more negative written review, while rational language coincides with a less negative written

review. However, it should be noted that this effect depends on the reputation expertise of the reviewer, as similar to the use of figurative speech in a negative review. To give examples, a negative review containing emotional speech written by a reviewer with a low reputation is more likely to be seen as less helpful and trustable. The same review written by a reviewer with high reputation is however more likely to be seen as helpful, due to the trusted (high) reputation of the reviewer (Sen and Lerman, 2007; Vermeulen and Seegers, 2009; Sparks and Browning, 2011). Furthermore, while considering a 5-point rating scale of stars, this study constructs very negative reviews by use of figurative and emotional speech, and a 1-star given rating, while less negative reviews are constructed by use of literal and rational speech, and a 2-star given rating.

To summarize, the effects of figurative and emotional speech in a negative review (in this study with the given review rating assumed as the valence of a review by construction of the variable degree of negativity) have a different impact on the booking likelihood of a reader and helpfulness of a review, depending on the reputation and trustworthiness of the reviewer (in this study assumed as the credibility of the reviewer by construction of the variable reviewer expertise). Additionally, it is argued by Lo and Yao (2019) that a review's valence and reviewer's credibility impacts the credibility of a review, where the latter is argued to impact the booking likelihood of a reader. This cross-dependence forms base for the interaction effect between the degree of negativity and reviewer expertise, and the construction of this study's conceptual framework and hypotheses (which are elaborated on in later sections).

2.2   Dependent variables

Most literature, which measures what readers think of a hotel after reading a review and a response, include either a dependent variable that measures a reader's evaluation score of the hotel, a reader's booking likelihood with the hotel or both. This study considers both variables as literature has also shown that a more favourable evaluation of a product or service does not necessarily imply that customers are more willing to buy a product (Casado-Díaz et al., 2020). Such finding can be explained as it is argued by Teas (1993) that if consumers are asked to evaluate a product or service, consumers personal preferences are of less importance. In contrast, if consumers are asked how likely they are to buy a product or service, their personal preferences are very important. For example, after reading a car manufacturer's response to a negative review, consumers may evaluate the manufacturer more positive. However, their booking likelihood may not change, as they do not like or consider the manufacturer's cars in the first place. Additionally, it has been shown by Volckner (2008) that consumers argue differently if they are asked about actions which affect their financial positioning (as booking a hotel). Therefore, the inclusion of these two dependent variables is needed to find out if the offering of compensation has a positive and significant effect on both a reader's booking likelihood and hotel evaluation score.

2.3 Conceptual framework

To determine when and how hotels should offer compensation in response to a negative online review, a framework based on existing literature is developed which contains independent and dependent variables, and the interaction effect. Thereby, these variables can be used to formulate different hypotheses, from which conclusions can be drawn on whether when and how hotels should offer compensation to negative online reviews.

From *Figure* 1, it can be seen that this study's conceptual framework is constructed by five different variables. This study tests for the relationship between the independent variable offering of compensation and the dependent variables booking likelihood and hotel evaluation. In this conceptual framework, that relationship is moderated by the interaction effect between the independent variables degree of negativity and reviewer expertise. It should be noted that, as Davidow (2000) argues against the importance of offering compensation, the outcome of this study's conceptual framework will cover a gap in literature about the effect of offering compensation in response to negative online reviews.
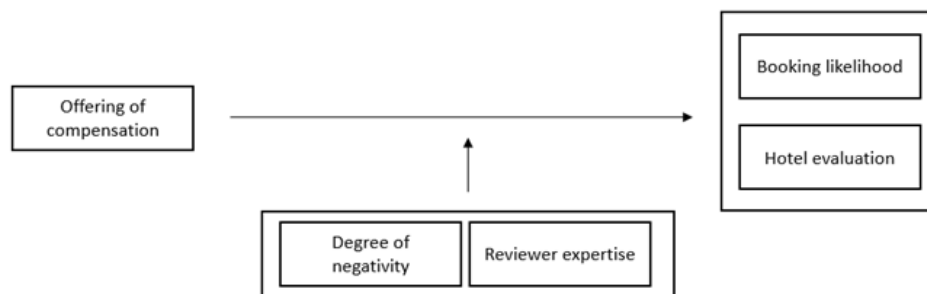
Figure 1: This study's conceptual framework.

2.4 Hypotheses development

The objective of this study is to determine when and how hotels should offer compensation in response to a negative online review. Hence, it is necessary to draw up different hypotheses, which may uncover components that lead to a better understanding of the relationship between the offering of compensation by a hotel and the booking likelihood and hotel evaluation of a reader.

The hypotheses are constructed as a combination of the interaction effect between the independent variables degree of negativity of a review and the expertise of a reviewer, the independent variable offering compensation by a hotel in response and the dependent variables booking likelihood and hotel evaluation score of a reader. As noted in *Section 2.1.2,* it is known that the effects of figurative and emotional speech in a negative review

have a different impact on the booking likelihood of a reader and helpfulness of a review, based on the reputation of the reviewer (it is assumed that these effects are also present in the interaction effect between the constructed variables 'Reviewer expertise' and 'Degree of negativity'). Additionally, it is argued that a review's valence and reviewer's credibility impacts the credibility of a review, where the latter is argued to impact the booking likelihood of a reader (a similar relationship is assumed for the hotel evaluation score of a reader). To the best knowledge of the writer, it is not known how this interaction effect influences the relationship between offering compensation and a reader's booking likelihood and hotel evaluation score. Therefore, in order to test how the relationship between the offering of compensation and a reader's booking likelihood and hotel evaluation score is influenced by this interaction effect, this study hypothesizes which kind of compensation companies should offer in order to obtain an as high as possible booking likelihood and hotel evaluation score of a reader given a certain negative review. Thereby, although constructing hypotheses in this way is unconventional, this study is able to give companies insight in when and how to offer compensation, even though literature on this topic is limited.

2.4.1 Very negative reviews written by reviewers with a high reputation

It is known that negative hotel reviews written by reviewers with a high reputation are seen as more helpful and trustable than negative reviews written by reviewers with a low reputation. Thereby, a review written by a reviewer with a high reputation has a larger impact on a reader's booking likelihood. It has also been shown that when figurative and emotional language is used in negative reviews and these reviews are written by reviewers with a high reputation, they are seen as more helpful and authentic in comparison to the same reviews of reviewers with a less trusted (low) reputation (Sen and Lerman, 2007; Vermeulen and Seegers, 2009; Sparks and Browning, 2011; Sparks and Bradley, 2014; Yin et al, 2020). Thereby, as the exact impact of the interaction effect (a very negative review written by a reviewer with a high reputation) on the relationship between offering compensation and booking likelihood is unknown, it is assumed that the impact of a very negative review (which in this study includes emotional and figurative speech) written by a reviewer with a high reputation on a reader's booking likelihood and hotel evaluation score is severe, due to the reputation of the reviewer and the intensity of the review.

It is hypothesized that if hotels offer full compensation in response to such review, this would lead to a significant higher booking likelihood and evaluation score in comparison to offering partial or no compensation at all. Such formulation is chosen, because very negative hotel reviews written by reviewers with a high reputation are assumed to have a very severe negative impact on hotel sales. Hence, it is assumed that hotels should offer full compensation in order to limit the impact of these reviews as much as possible, as Piehler et al. (2019) argues that the offered compensation should match the intensity of the customer complaint. Additionally, offering partial compensation is hypothesized to lead to a significantly higher booking likelihood and hotel evaluation score of a reader in comparison to offering no compensation. Such formulation is based on Mount and Matilla (2000), who argue that even the smallest offering (as for example a gift card) will lead to favourable customer

satisfaction in comparison to offering no compensation. Hence, it is hypothesized that offering partial compensation is beneficial to offering no compensation even though the offering of partial compensation is assumed to not match the severity of the customer complaint (a very negative review) as argued for by Piehler et al. (2019).

To be able to determine how the relationship between the offering of compensation and booking likelihood and evaluation score is impacted by a very negative review written by a reviewer with a high reputation, the following hypothesis is considered:

Hypothesis 1: Considering a very negative review written by a reviewer with a high reputation, the offering of full compensation leads to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while the offering of partial compensation leads to significantly higher booking likelihood and hotel evaluation score in comparison to not offering any compensation.

2.4.2   Less negative reviews written by reviewers with a high reputation

Likewise to the first hypothesis, it is unknown how the interaction effect between less negative reviews and reviewers with a high reputation impacts the relationship between the offering of compensation and the reader's booking likelihood and hotel evaluation score. As this study has constructed less negative reviews as consisting of literal and rational speech, and a 2-star rating, it is assumed that the impact of a less negative review on the booking likelihood and hotel evaluation score of a reader is less severe in comparison to very negative reviews. Based on results by Piehler et al. (2019), it is assumed that offering partial compensation would match the severity of the customer complaint, while offering full compensation is assumed to be overcompensation. Such reasoning is supported by recent research into the offering of overcompensation to a customer complaint in general by Gelbrich and Roschk (2011). They argue that overcompensation only slightly increases overall satisfaction in comparison to compensation which matches the intensity of the complaint. To give an applied example from within the hotel industry, if a customer is fully compensated for a hotel room, because the dirty bed sheets were not changed, overall customer satisfaction will only be slightly higher in comparison to overall satisfaction when a fifty percent discount coupon for a future stay is offered (Noone and Lee, 2011). It should be noted that Noone and Lee (2011) only look into the direct effect of offering compensation and overall customer satisfaction and do not consider any similar interaction effect as the one in this study.

Based on these results, it has been chosen to hypothesize that the offering of partial compensation is equally beneficial to the offering of full compensation when considering responses to less negative reviews written by highly regarded of reviewers. Hence, it is assumed that in this setting, the offering of full compensation would be considered as overcompensating for a not too severe complaint (as a less negative review). Thereby, it is assumed that offering partial compensation in this setting does matches the magnitude of the complaint and will lead to a significant higher booking likelihood and evaluation score in comparison to offering no compensation at all.

To be able to determine how the relationship between the offering of compensation and booking likelihood and evaluation score is impacted by a less negative review written by a reviewer with a high reputation, the following hypothesis is considered:

Hypothesis 2: Considering a less negative review written by a reviewer with a high reputation, the offering of full compensation does not lead to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while offering both levels of compensation does lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering of any compensation.

2.4.3   Very negative reviews written by reviewers with a low reputation

It is known that a negative review written by a reviewer with a low reputation has a lower impact on the booking likelihood of a reader and is considered as less helpful and trustable in comparison to the same review by a reviewer with a high reputation. It has also been shown that readers are less convinced of the authenticity of the review, while the impact of this interaction effect on the relationship between offering compensation and booking likelihood and hotel evaluation score of a reader is unknown (Sen and Lerman, 2007; Vermeulen and Seegers, 2009; Sparks and Browning, 2011; Sparks and Bradley, 2014; Yin et al, 2020). Thereby, it is assumed that the impact of a very negative review written by a reviewer with a low reputation on a reader's booking likelihood and hotel evaluation score is very limited (or low), as Xu (2014) argues that readers tend to find the reputation of the writer more important, when reading a very negative review in comparison to reading a less negative review. Xu (2014) also argues that negative hotel reviews written by a reviewer with a low reputation may be biased in the eyes of readers, as readers may believe that these kind of reviews are not representative for the true performance of a hotel and that the reviewer has possibly only created an account to bash the hotel online.

Similar results are found by Chua and Banerjee (2015), who argue that readers are more likely to ignore such kind of reviews, because these reviews are less helpful, have less impact and are seen as less trustable to readers in comparison to the same reviews written by reviewers with a high reputation. Chua and Banerjee (2015) explain this finding by giving an example of the reviewer ranking system of a booking website. Readers use the ranking system of a booking website (as TripAdvisor or Booking.com) as an indicator to reckon if a review writer (and thereby the review) is trustful or not. Reviewers, that write many reviews and receive good feedback from other users, accumulate a lot of points and are, therefore, seen as more trustworthy than others. Consequently, their reviews are seen as more trustworthy and helpful in comparison to reviews of reviewers with a low reputation, which are seen as less trustworthy and helpful, and are, therefore, more likely to be ignored.

Another explanation to why readers may ignore these reviews is given by Sen and Lerman (2007), who argue that readers of reviews of hedonic (or luxury) products are more likely to discount such reviews, because they have higher expectations of hedonic products. To give an example, a very negative review about a hedonic product or service, as a stay in a hotel, is more likely to be discounted, because readers have higher expectations

of hedonic products than utilitarian (basic needs as food and water) ones. It is argued that readers argue counter-intuitively and look for reasons why they can discount such review. Examples of such reasoning could be the lack of reputation of the review writer or the belief that the review writer is exaggerating, a finding in line with Xu (2014). However, it should be noted that, when reviewers have a high reputation, readers are more likely to be convinced of the authenticity of a negative review (Sen and Lerman, 2007; Vermeulen and Seegers, 2009; Sparks and Browning, 2011).

When considering these results by literature, it seems that readers do not think high of a very negative reviews written by reviewers with a low reputation and it is assumed that they are likely to ignore them. Thereby, it can be argued that offering full or partial compensation can be regarded as overcompensating, as the impact of these kinds of reviews on readers is assumed to be limited. Hence, while companies may try to limit the impact of a very negative review written by a reviewer with a low reputation, in offering full or partial compensation they are assumed to overcompensate, as readers are likely to ignore these kinds of reviews. Additionally, when considering that companies cannot offer compensation to each negative review due to financial limitations (Piehler et al., 2019), it makes sense for companies to not offer compensations in response to reviews which are likely to be ignored by readers. However, it is suggested by Yin et al. (2020) that readers ignore these kinds of reviews only if they have access to other reviews (as less negative or positive ones) about the same hotel, but proof of this relationship, when only very negative reviews written by reviewers with a low reputation are available, is yet to be established. As this study by construction isolates reviews based on the valence of a review, the outcome of this hypothesis adds to fill a gap in literature. Moreover, it is hypothesized that offering no compensation at all in response to very negative reviews written by reviewers with a low reputation would lead to the same booking likelihood and evaluation score as offering full or partial compensation.

To be able to determine how the relationship between the offering of compensation and booking likelihood and evaluation score is impacted by a very negative review written by a reviewer with a low reputation, the following hypothesis is considered:

Hypothesis 3: Considering a very negative review written by a reviewer with a low reputation, the offering of full or partial compensation does not lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering any compensation.

2.4.4 Less negative reviews written by reviewers with a low reputation

Similar to the previous hypotheses, it is unknown how the interaction effect between less negative reviews and reviewers with a low reputation impacts the relationship between the offering of compensation and the reader's booking likelihood and hotel evaluation score. As this study has constructed less negative reviews as consisting of literal and rational speech, and a 2-star rating, it is assumed that the impact of a less negative review on the booking likelihood and hotel evaluation score of a reader is less severe in comparison to very negative reviews. However, as very negative reviews written by reviewers with a low reputation are more likely to be ignored and,

therefore, are assumed to only have a very limited influence on the booking likelihood of a reader, less negative reviews written by reviewers with a low reputation are assumed to have a bigger impact on the booking likelihood of a reader.

Results of literature which state that very negative reviews written by a reviewer with a low reputation are likely to be ignored by readers, often note that such relation does not exist for less negative reviews. While very negative reviews by these reviewers are sometimes seen as exaggerated, less negative reviews by these reviewers are seen as more genuine (Xu, 2014; Yin et al., 2020). Such finding can be explained as less negative reviews often include less angry language and are seen as more rational and are, therefore, more helpful to readers (Craciun and Moore, 2019; Yin et al., 2020). Readers are also more likely to trust reviewers with a low reputation who write a less negative review than ones who write very negative reviews, because readers are less likely to believe that these reviewers made an account to purely bash a hotel. Thereby, less negative reviews written by these reviewers are more likely to be seen as a true reflection of the performance of a hotel than very negative reviews written by these kinds of reviewers (Xu, 2014).

Other literature argues against the offering of compensation in response to less negative reviews written by reviewers with a low reputation. Initially, Piehler et al. (2019) argues in favour of responding by offering compensation, but thereafter notes that it may not be financially beneficial for companies to offer compensation in response to all negative reviews. To give an example, when a company is cash-strapped, it should only allocate resources to the most negative and influential reviews (where it should be noted that Piehler et al. (2019) does not includes an interaction effect similar to the one in this study). Hence, it is assumed that offering compensation to less negative reviews written by reviewers with a low reputation would be of least importance. Additionally, it is suggested that the influence of less negative reviews written by these reviewers on readers is unsignificant, because readers consider less negative reviews as very similar to a neutral ones. Hence, it is argued that the influence of the less negatively written review is negligible, especially when written by a reviewer with a low reputation (Ye, 2009).

This study hypothesizes that hotels should offer compensation to less negative reviews written by reviewers with a low reputation, because literature has shown that, although the influence of such reviews is very limited and companies should only allocate financial resources to the most negative reviews and most influential reviewers (Ye, 2009; Piehler et al., 2019), such reviews are also seen as more genuine, trustable and, therefore, more helpful to readers and are more likely to represent the true performance of a hotel in comparison to very negative ones. Therefore, it is assumed that correctly responding (by offering compensation) to these reviews is more important than correctly responding to very negative reviews by written reviewers with a low reputation, because the latter reviews are more likely to be ignored by readers. Hence, it is hypothesized that, as the offering of compensation should match the intensity of the customer complaint (Piehler et al., 2019), offering full compensation to a less negative review written by a reviewer with a low reputation would be overcompensating in comparison to offering partial compensation, while offering no compensation is assumed to lead to a lower

booking likelihood and hotel evaluation score of a reader, as these reviews are seen as genuine, trustable and helpful to readers, and the review itself (while not very negative) is negative.

To be able to determine how the relationship between the offering of compensation and booking likelihood and evaluation score is impacted by a less negative review written by a reviewer with a low reputation, the following hypothesis is considered:

Hypothesis 4: Considering a less negative review written by a reviewer with a low reputation, the offering of full compensation does not lead to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while offering both levels of compensation does lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering of any compensation.

2.5   Overview of the hypotheses

An overview of the hypotheses is given in *Table 2*. In summary, it is firstly hypothesized that offering full compensation will lead to a statistically significant higher booking likelihood and hotel evaluation score in comparison to partial compensation and no compensation offered, when considering a very negative review written by a reviewer with a high reputation. Additionally, offering of partial compensation will lead to significant higher results when compared to not offering any compensation at all. Such formulation is chosen, because it is assumed that a very negative review written by a reviewer with a high reputation will have a severe negative impact (denoted by '--' when considering the Booking likelihood / Hotel evaluation score in *Table 2*) on the booking likelihood of a reader (and by assumption on the hotel evaluation score). Additionally, the review reputation of the reviewer ensures that the review is seen as more helpful and trustable than a less negative review (denoted by '+' when considering the Helpfulness / Trustworthiness in *Table 2*). Hence, it is hypothesized that full compensation should limit the negative impact of these reviews significantly better than offering partial or no compensation. Additionally, although it is assumed that offering partial compensation does not match the magnitude of the customer complaint, offering partial compensation should limit the impact of these reviews better than offering no compensation, because offering any compensation has been shown to increase customer satisfaction in comparison to not offering any compensation at all.

Secondly, it is hypothesized that offering of full and partial compensation will lead to equally statistically significant higher results in comparison to not offering any compensation when considering a less negative review written by a reviewer with a high reputation. Such formulation is chosen, because it is assumed that a less negative review written by a reviewer with a low reputation has a less severe negative impact (denoted by '-' when considering the Booking likelihood / Hotel evaluation score in *Table 2*) on the booking likelihood of a reader than a very negative review written by the same reviewer. It is also assumed that offering full compensation would be overcompensating, due to the magnitude of the customer complaint (a less negative review). In addition, offering partial compensation is assumed to match the magnitude of the customer complaint. Offering no compensation is assumed to lead to lower results, as the level of expertise of the reviewer

is high and the review itself (while not very negative) is negative. It should be noted that, as the reviewer has a high reputation in both cases, the helpfulness and trustworthiness in the first and second hypotheses is rated as equal (as denoted by '-' when considering the Helpfulness / Trustworthiness in both cases in *Table 2*).

	High reputation	Low reputation
Very negative	Booking likelihood / Hotel evaluation score: -- Helpfulness / Trustworthiness: + H1: Full > Partial > No	Booking likelihood / Hotel evaluation score: +/- Helpfulness / Trustworthiness: -- H3: Full = Partial = No
Less negative	Booking likelihood / Hotel evaluation score: - Helpfulness / Trustworthiness: + H2: Full = Partial > No	Booking likelihood / Hotel evaluation score: - Helpfulness / Trustworthiness: +/- H4: Full = Partial > No

Table 2: An overview of hypotheses 1 till 4 is given. It should be noted that 'Full', 'Partial' or 'No' corresponds to the level of offered compensation in a hotel response. Additionally, the indicator '>' denotes an outperformance in higher booking likelihood and hotel evaluation score by a certain level offered compensation in comparison to another level of compensation, while '=' denotes an equal performance in comparison to a different level of offered compensation. For each hypothesis, the effects based on literature on booking likelihood / hotel evaluation score of a reader (where '--' denotes a severe negative impact, '-' denotes a less severe impact and '+/-' denotes a neither negative or positive impact on the booking likelihood / hotel evaluation score of a reader) and the interaction effect between helpfulness and trustworthiness of the review and reviewer (where '+' denotes a helpful/trustable review, '+/-' denotes a neither helpful/trustable or unhelpful/untrustable review and '--' denotes a very unhelpful/untrustable review) is given.

Thirdly, it is hypothesized that it doesn't matter which level of compensation is offered, the booking likelihood and hotel evaluation score of a reader will not be significantly different, when considering a very negative review written by a reviewer with a low reputation. Such formulation is chosen, because readers are more likely to ignore these reviews (denoted by '--' when considering the Helpfulness / Trustworthiness in *Table 2*) and, therefore, the negative impact of these reviews on the booking likelihood of a reader is very limited (denoted by '+/-' when considering the Booking likelihood / Hotel evaluation score in *Table 2*). Hence, it is hypothesized that there is no significant difference in booking likelihood and hotel evaluation score of a reader between offering full, partial or no compensation.

Fourthly and lastly, when considering a less negative review written by a reviewer with a low reputation, the offering of full and partial is hypothesized to lead to equally statically significant higher results in comparison to offering no compensation at all. Such formulation is chosen, because these reviews, although written by reviewers with a low reputation, are more likely to be seen as a true reflection of the hotel's performance (denoted by '+/-' when considering the Helpfulness / Trustworthiness in *Table 2*) and have a larger negative impact on the booking likelihood of a reader (denoted by '-' when considering the Booking likelihood / Hotel evaluation score in *Table 2*) than very negative reviews written by the same reviewer. It assumed that offering full compensation would be overcompensating due to the magnitude of the customer complaint (a less negative review). Additionally, it is assumed that offering no compensation would lead to a lower booking likelihood and hotel evaluation score of a reader in comparison to offering partial compensation, as these reviews are seen as genuine and helpful to readers, and the review itself (while not very negative) is negative. Furthermore, it should be noted that it is assumed that offering compensation to these reviews is more important than offering compensation to very negative reviews written by reviewers with a low reputation, because the latter are more likely to be ignored.

3   Methodology

3.1   Research design

This study uses an experimental design to empirically test for the hypotheses within the conceptual framework, by implementing a survey. Such choice is made as this study considers many independent and dependent variables. Surveys enable for consistent and comparable results between participants, because the options for participants (who represent the readers) to respond are limited and standardized in comparison to other data collection methods as for example interviews or focus groups. Besides, the use of a survey enables this study to be time and cost efficient, while being able to collect data in the mid of the COVID-crisis. Surveys also enable for anonymous data collection, which ensures that participants may not feel the necessity to respond accordingly to peer pressure, which increases the accuracy of this study's results (Mathers, 1998).

By using a survey, this study tries to uncover new elements which may contribute to the understanding of the relationship between offering compensation by a hotel and the booking likelihood and evaluation score of a reader. Thus, this study looks into the effects of the independent variable offering of compensation on the two dependent variables (i.e., booking likelihood and hotel evaluation score), while considering the moderating interaction effect between the variables degree of negativity and reviewer expertise. Hence, this study is constructed as a 2 (degree of negativity: less or very) x 2 (reputation of the reviewer: high or low) x 3 (offering of compensation: no, partial or full) between-subjects design containing 12 experimental conditions, all of which can be seen in *Table 3*:

#	Degree of negativity	Reviewer expertise	Compensation
1	Very negative	High	Full
2	Very negative	High	Partial
3	Very negative	High	None
4	Very negative	Low	Full
5	Very negative	Low	Partial
6	Very negative	Low	None
7	Less negative	High	Full
8	Less negative	High	Partial
9	Less negative	High	None
10	Less negative	Low	Full
11	Less negative	Low	Partial
12	Less negative	Low	None

Table 3: An overview of all experimental conditions considered in the main experiment.

However, as literature on this experimental design is limited, this study made assumptions to construct a conceptual framework. It is assumed that the distinction between less and very negative reviews, and reviewers with a high or low reputation is constructed correctly. Thereby, it could be argued that these made assumptions, if not tested for rightfulness, may bias the outcome of this study. Hence, it is necessary to test if such made assumptions hold in an experimental setting by using a pretest. The pretest will test if participants can distinct

between the different levels within the categorical independent variables. To give an example, it is tested if participants of this pretest can distinct very negative reviews from less negative reviews. Hence, participants are asked to read a negative review and rate the degree of negativity of the review by use of a Likert-scale. If participants are able to successfully distinct between for example different types of negative reviews, this implies that the constructed distinction between very and less negative reviews was done correctly. However, it should be noted that this only implies that participants are able to distinct very negative reviews (based on figurative and emotional speech and a 1-star review rating) from less negative reviews (based on literal and rational speech and a 2-star rating). Moreover, the rightfulness of this study's conceptual framework depends on the outcome of the pretest. If results are positive, this study's assumption and distinctions behind the construction of the conceptual framework were made correctly and hypotheses testing will continue as given. However, if the results of the pretest are negative, the conceptual framework will be altered for accordingly.

3.2   Continuity and consistency across research

To make sure that the results of each filled in survey can be compared to another, the 12 experimental conditions are constructed to be as similar as possible. Therefore, all negative reviews, reviewer profiles and responses of hotels are constructed to resemble their real life counterparts as closely as possible.

3.2.1   Construction of negative reviews and reviewer profiles

In the main experiment, two negative reviews are constructed by incorporating a certain degree of negativity and two reviewer profiles are constructed by incorporating a certain level of reviewer expertise. Which reviews and reviewer profiles are used depends on the outcome of the pretest. The following sections will give examples of how these reviews and profiles are constructed. All the considered reviews and review profiles in the pretest and the main experiment of this study are included in Appendix A.

Degree of negativity

The distinction between very and less negative reviews was made by considering very negative reviews to be written by the use of figurative and emotional speech and a 1-star awarded review rating, while less negative reviews award a 2-star rating and are written in literal and rational speech. While the distinction made between a 1-star and 2-star review rating (on a scale of 1 till 5) is clear, the difference between figurative and emotional speech and literal and rational speech is less comprehensive. To illustrate the difference between very and less negative reviews within this study, two reviews from booking website TripAdvisor are considered in *Table 4*.

As part of making the distinction between 1-star and 2-star review ratings, it was noted that a 2-star rated review often includes a positive note, while a 1-star rated review does not. From the examples in *Table 4*, it can be seen that while the less negative review is written negatively, it does include some positive notes as the reviewer names the location and view of the room as positives. Yet, the very negative review does not include any positive notes. Moreover, the difference between figurative and literal speech is clearly illustrated as the writer of the

very negative review uses a lot of emotional and figurative speech. Examples include '*I cannot convey strongly enough how disgusting this place is*' and '*due to sanitation conditions similar to those of a homeless crack head.*'. Contrarily, the writer of the less negative review uses more literal and rational speech, as can be seen from '*the view was great, but the bed was very uncomfortable*' and '*the chair height was not optimal for working at the desk pictured on the website*'.

Very negative: 1-star, figurative and emotional speech	Less negative: 2-star, literal and rational speech
★☆ ☆ ☆ ☆ *"There was poo in the kettle…* *Do not stay in this ABSOLUTE dump of a hotel, and I use the term hotel loosely. I cannot convey strongly enough how disgusting this place is. Blood stained headboards that have clearly been up since the world war (the first one), rude staff, windows that won't close, no hot water, broken furniture, dirty utensils, broken light fixings and actual poo in the kettle. Pretty sure I'm going to end up with some sort of rash / disease due to sanitation conditions similar to those of a homeless crack head. In summary…this place is a complete hole."*	★★☆ ☆ ☆ *"Great location but uncomfortable stay* *Stayed here for a few nights. Found the location to be very convenient to access Chinatown and surrounding areas. We paid for a breakfast and cocktail package. The breakfast was awful, and we opted on not using our breakfast credits while we were there after two bad experiences. In addition, the premade cocktails were also bad. As for the room the view was great, but the bed was very uncomfortable. In addition, the chair height was not optimal for working at the desk pictured on the website."*

Table 4: An example of a very negative hotel review is given on the left, while an example of a less negative hotel review is given on the right.

Furthermore, it can be noticed from the TripAdvisor examples in *Table 4* that, while both reviews are written by hotel visitors, the covered topics differ heavily. Therefore, it has been chosen to fix the number of covered topics in the negative reviews to four. Independently of degree of negativity, reviewer expertise or the compensation offered, all considered reviews in the pretest and main experiment (less or very negative) will include a statement made about the hotel room, location, service and price, because it is argued by Hu (2019) that these are the most discussed topics in online hotel reviews.

Reviewer expertise

The expertise of a reviewer is constructed as reviewers with a high or low reputation, while the reviewer profiles in this study are based on the reviewer profiles of booking website TripAdvisor. Reviewers with a high reputation have a verified account with a profile picture of a human, are described as 'Travel expert' and have written a lot

of contributions (naming for reviews on TripAdvisor) and their reviews have received a lot of helpful votes. Contrarily, reviewers with a low reputation have an anonymous account without a profile picture, are described as 'Travel amateur' and have not written a lot of contributions and their reviews have not received a lot of helpful votes. To illustrate the difference between reviewers with a high or low reputation, two imaginary profiles are considered in *Table 5*:

High reputation: verified account, profile picture, 'Travel expert' and high number of awarded points	Low reputation: anonymous profile, 'Travel amateur' and low number of awarded points
_SaraWang ✔ *Travel expert • 1986 contributions • 4637 helpful votes*	S _SaraWang *Travel amateur • 1 contribution • 0 helpful votes*

Table 5: An example of a reviewer with a high reputation is given on the left, while an example of a reviewer with a low reputation is given on the right.

From *Table 5*, it can be seen that the difference between a reviewer with a high reputation and a low reputation is easy to notice. While the username of the reviewer is identical over the two cases, the profile of the reviewer with a high reputation contains eye catching features as a profile picture and a verified badge. Additionally, it can be seen that the reviewer is denoted as a 'Travel expert', which coincides with a large number of contributions and helpful votes. Contrarily, the profile of the reviewer with a low reputation is less catchy, as it only features a generic and anonymous profile picture and misses the blue verified badge. In addition, it can be noticed that the reviewer has only written one contribution and received zero helpful votes. These scores coincide with the status of the reviewer, who is denoted as 'Travel amateur'.

3.2.2   Construction of hotel responses

The hotel responses in this study are constructed to be as similar as possible. A hotel's response in one experimental condition should ideally only differ from another by the offered compensation. Hence, the objective is to construct hotel responses as such that the effect of the different levels of offered compensation is not affected by other elements within a review. To give an example of such, literature shows that two elements are considered to be important by readers when reading a response to a negative review, namely the writing style and the timing of the response (Park and Allen, 2013; Sparks and Bradley, 2014). Thereby, a slow response and offer of full compensation could potentially be outperformed by a quick response and offer of partial compensation. However, it is unsure what the real effect of the level of offering compensation in this example is, as it could be either that the timing of the response or the offering of compensation caused for the outperformance. Hence, in order to obtain valid results, the timing of the response is kept constant in this study.

Writing style

Literature shows that a response to a negative review should incorporate some standard components, as Sparks (2014) argues that a certain typology should be present in those responses. This typology, known as the 'Triple A' typology, accounts for three main components when responding to negative online reviews, namely acknowledgment of the review, accounting for the customer complaint and taking action to resolve the relationship with the customer. To give examples, it is argued that hotel responses should include a form of acknowledgment towards the review writer, as for example an offered apology. Additionally, responses should take account for the statement the customer is making, which can be for example an explanation or justification for a customer complaint. Further, responses should address how the customer's complaint is going to be handled and what actions are taken to prevent such flaw from happening again in the future. In her research, Sparks (2014) has given some examples of sentences based on the Tripe A typology. Based on these examples, a hotel response is constructed, which will be used in the main experiment of this study and is considered in *Table 6*:

Dear Sara,

Thank you for taking the time to write a review of your recent stay at our hotel. We greatly appreciate your comments.

It was concerning to read about the issues you encountered during your recent stay and I would like to apologize for the inconvenience it caused. I have personally inspected your room and it seems that we dropped the ball in a few areas during your visit.

Unfortunately during your stay, the city council had ordered a partial lockdown due to the COVID-19 pandemic. As we informed you, this implied that we were working with less than half of our staff. I have met with our whole staff to ensure that any further inconvenience will be minimized in the future.

Although we try to create the best experience possible for our hotel guests during a global pandemic, we endeavour to meet the room requirements of all our guests at all times. Thereby, we would like to fully compensate you for your stay at our hotel and hope that you will visit us again in the future.

I will contact you personally to discuss matters forward.

We hope your next stay at our hotel will be a pleasant journey and look forward to welcoming you back again.

Kind regards,

John Watson
Guest Relations Manager

Table 6: The considered hotel response used within the main experiment of this study. The response is constructed accordingly to the Triple A typology of Sparks (2014).

From *Table 6*, it can be noticed that the response by the hotel consists of the three main components of the Triple A typology. First, the red lines in the example form the acknowledgement of the review by including an appreciation of the taken time to write a review, while offering an apology. Second, the blue lines denote the part of the response were the hotel managers takes account for the customer complaint by explaining what caused the inconvenience. Third and last, the manager describes which actions are taken to prevent such from happening again in the future, as indicated by the green lines. Additionally, the content of the response is written to be honest, thorough , adequate and professional, while also being as friendly and informal as possible (Sparks, 2014). Furthermore, it should be noted that, while this response offers full compensation in response, which compensation is offered in the main experiment this study depends on which experimental condition is considered. All three considered responses considered in this study are included in Appendix A.

Timing of the response

Another important variable in responding to negative online reviews (or customer complaints in general) is the response time (Mount and Mattila, 2003, Sparks et al., 2016; Li et al., 2017). Overall customer satisfaction increases when a response is given, but increases even more when the given response is given within two business days (Sparks et al., 2016). Additionally, literature shows that the content of a response has a larger impact on customer satisfaction if the response is given quickly (Sparks and Bradley, 2014). To give an example, quickly apologizing and taking corrective actions tends to lead to higher overall customer satisfaction than slowly apologizing and taking corrective actions. Therefore, it is argued that the timing of a response is an important element in responding to negative online reviews and part of the Triple A typology (Sparks and Bradley, 2014; Sparks et al., 2016). Furthermore, it is noted that the timing of the response does not have a significant impact when the content of the response is not considered to be sufficient in comparison to the magnitude of the complaint described in the review. To be more specific, if for example a quick response of a hotel only acknowledges the customer's complaint and does not take sufficient action to solve the complaint, the review writer may feel that the complaint is not taken serious enough and the benefit of quick response is lost (Sparks and Bradley, 2014). Thereby, as this study is keen to find out what the true effect of offering different levels of compensation is, all hotel responses in this study are given within one business day.

3.3   Pretest

The objective of the pretest is to ensure that the made distinctions within the independent variables in study are done correctly. Hence, the pretest tests for the rightfulness of all made distinctions. Each participant will be presented with six different reviews and six different reviewers. For continuity and consistency, each of the three less negative reviews are constructed to have a very negative counterpart, where a combination of both includes the same topics and only differs in the degree of negativity. The participant will be asked to rate the degree of negativity of a review and the expertise of the reviewer. All presented questions to the participants of the pretest are summarized in *Table 6*. In addition to these questions, four different control questions are included in the pretest. Participants are asked to fill in their gender, age, current occupation and highest obtained degree. An overview of the considered reviews and reviewers as the control questions are included in Appendix A.

#	Independent variable	Question	Measurement	Answer possibilities	Outcome
1	Reviewer expertise	How would you rate the level of expertise of this reviewer?	Likert scale	1) Very low 2) Somewhat low 3) Neutral 4) Somewhat high 5) Very high	Shows if reviewers with a high reputation (verified profile, profile picture, 'Travel expert' and a high number of awarded points) can be distinct from reviewers with a low reputation (anonymous profile, 'Travel amateur' and a low number of awarded points).
2	Degree of negativity	How negative or positive do you think this review is?	Likert scale	1) Very negative 2) Somewhat negative 3) Neutral 4) Somewhat positive 5) Very positive	Shows if very negative reviews (figurative, emotional speech and a 1-star rating) can be distinct from less negative reviews (literal, rational speech and a 2-star rating).

Table 7: An overview of all research questions included in the pretest. It should be noted that the number '#' denotes the order of presented questions.

It can be noted from *Table 7*, that the answer possibilities of the questions include a five point Likert scale. Such construction is chosen for as Joshi (2015) argues that Likert scales are well equipped to test for the attitude (or the opinion) of participants by giving participants a full range of answers (i.e., strongly disagree, agree, neutral, agree and strongly agree) to choose from. Additionally, the use of Likert scales ensures for consistent and easy to analyse results. However, in order to use a Likert scale, at least five answer possibilities should be available to the participant. In order to test how the participant rates the level of expertise of the reviewer, five answer possibilities are considered, ranging from 'very low' till 'very high'. Additionally, in order to test how the participant rates the degree of negativity of the review, five answer possibilities are considered, ranging from 'very negative' till 'very positive'. It should be noted that it has been chosen to formulate 'less negative' as 'somewhat negative', because it assumed that participants will be able to easier distinct 'very negative' from 'somewhat negative' in comparison to 'very negative' from 'less negative'.

3.4   Main experiment

The objective of the main experiment is to test for the hypotheses within the conceptual framework. Each participant will be presented with one of the 12 experimental conditions. To give an example, a condition can be a very negative review written by a reviewer with a high reputation and a response by a hotel offering full compensation in response. It should be noted that the presented reviews and reviewer profiles in the main experiment were selected out of the reviews and profiles considered in the pretest and are included in Appendix A.

Regardless of condition, all participants are asked to answer to the same ten questions. All presented questions of the main experiment are summarized in *Table 8*. In addition to these questions, the same four control questions as within the pretest are considered. Additionally, participants are also asked to rate their concern about COVID-19, as the simulated response by the hotel includes a reference to the virus and a participants concern about the virus may influence their booking likelihood. An overview of the considered control questions is included in Appendix A.

#	Variable	Question	Measurement	Answer possibilities
1	Booking likelihood	How likely are you to book a stay at this hotel?	Likert scale	1) Very unlikely 2) Unlikely 3) Neither unlikely nor likely 4) Likely 5) Very likely
2	Evaluation score	How would you rate this hotel? Please note that a higher number denotes a higher rating score.	Ordinal	Number range of 0 till 10
3	Recommendation	How likely are you to recommend this hotel to friends and/or family?	Likert scale	1) Very unlikely 2) Unlikely 3) Neither unlikely nor likely 4) Likely 5) Very likely
4	Response rate	How would you rate the response by the hotel given the review?	Likert scale	1) Very ungenerous 2) Ungenerous 3) Neither ungenerous or generous 4) Generous 5) Very generous
5	Response satisfaction	If you had written this review, how satisfied would you be with the response by the hotel?	Likert scale	1) Very dissatisfied 2) Dissatisfied 3) Neither dissatisfied nor satisfied 4) Satisfied 5) Very satisfied
6	Offering compensation	What kind of compensation is offered to Sara?	Ordinal	1) None 2) Partial compensation 3) Full compensation
7	Reviewer expertise	How would you rate Sara's level of expertise as a reviewer?	Likert scale	1) Very low 2) Low 3) Neither low or high

#				4) High
				5) Very high
8	Reviewer trustworthiness	How would you rate Sara's trustworthiness as a reviewer?	Likert scale	1) Very untrustable 2) Untrustworthy 3) Neither untrustable nor trustable 4) Trustable 5) Very trustable
9	Degree of negativity	How negative or positive do you think this review is?	Likert scale	1) Very negative 2) Somewhat negative 3) Neither negative or positive 4) Somewhat positive 5) Very positive
10	Review helpfulness	How would you rate the helpfulness of Sara's review?	Likert scale	1) Very unhelpful 2) Somewhat unhelpful 3) Neither unhelpful nor helpful 4) Somewhat helpful 5) Very helpful

Table 8: An overview of all research questions included in the main experiment. It should be noted that the number '#' denotes the order of presented questions.

This study's main experiment is constructed as fifteen different questions, which are asked to a participant after he or she has read a review and response within an experimental condition. From the ten main questions in *Table 8*, the first three questions account directly for the outcome of the conceptual framework. Beforehand, participants are asked to imagine that he or she is considering the hotel in the experimental condition for a holiday stay and is reading its reviews. Moreover, as this study's outcome depends on the booking likelihood and evaluation score of participants, it is necessary that participants give answers as if they are in the process of deciding which hotel to go to on holiday and truly measure the effect of their perception of a negative review and hotel response on their booking likelihood and hotel evaluation score. The participant is directly asked how likely he or she is to book a stay at this hotel after reading the review and the hotel's response. However, it is argued by Lien (2015) that a consumer's booking likelihood is also determined by how likely the consumer is to recommend a product or service to friends and/or family and, therefore, question 3 is included.

In questions 4 and 5, the participant is asked about his or her opinion about the offered compensation by the hotel, in order to understand how the participant feels about the offered compensation. These questions coincide with question 6, which checks if the participant can correctly identify which level of compensation was offered by the hotel as part of a manipulation check. Furthermore, while the pretest shows if participants are able to recognize distinctions made between the independent variables, question 9 is also included as part of

the manipulation check in order to check if participants are able to defer between very and less negative reviews. Such construction is necessary, because it would be false to assume that participants of the main experiment can automatically distinct very negative from less negative reviews based on results by the participants of the pretest. Hence, a manipulation check in the main experiment is needed to ensure that participants have actually read and understood the given review and response.

In questions 7, 8 and 10, the participant is asked to rate the expertise (as the final part within the manipulation check) and trustworthiness of the reviewer, and the helpfulness of the review, as it was shown by literature that different levels of degree of negativity (very or less) and reviewer expertise (high or low) are seen as more or less trustable or helpful to a reader (Craciun and Moore, 2019; Yin et al., 2020). As these results are widely used in order to draw up the hypotheses of this study, these questions are included to gain further insight in why a participant rates a hotel high or low after reading a review and a certain response.

4   Results of the pretest

Data for both surveys was anonymously collected by using Qualtrics online surveys. Both surveys were distributed among friends and acquaintances of the writer, where participants of the pretest survey were not asked to fill in the survey of the main experiment to ensure that both surveys were independently collected of each other. The pretest' survey was ran on June 12th 2021 and was filled in 51 times before collection was stopped. The survey of the main experiment ran between the 16th and 22nd of June 2021 and was fully filled in 494 times before collection was stopped. The survey was constructed as a mobile and pc version from which the results were combined to form one data set. All partial filled in surveys were excluded from the survey response. In addition, all survey responses to the main experiment which did not correctly recognized which compensation was offered in response were also excluded from the survey response. For analysis, it should be noted that it is assumed that Likert scales have an interpretable mean and are normally distributed if the number of filled in surveys for each experimental condition is bigger than 30 (N > 30).

4.1   Control variables

Variable	Frequency (N = 51)	Percent	Cum. percent
Gender			
- Male	26	51.0	51.0
- Female	25	49.0	100.0
Age			
- 16/25	21	41.2	41.2
- 26/35	30	58.8	100.0
Occupation			
- Employed full time	39	76.5	76.5
- Employed part time	3	5.9	82.4
- Unemployed looking for work	1	2.0	84.3
- Student	8	15.7	100.0
Degree			
- High school graduate	2	3.9	3.9
- MBO	4	7.8	11.8
- Bachelor	30	58.8	70.6
- Master	15	29.4	100.0

Table 9: An overview of the descriptive statistics of the control variables of the pretest survey. Please note: Gender (1 = Male, 2 = Female, 3 = Non-binary / third gender, 4 = Preferred not to say), Age (1 = Younger than 16 years old, 2 = 16 till 25, 3 = 26 till 35, 4 = 36 till 45, 5 = 46 till 55, 6 = 56 till 65, 7 = Older than 65 years old), Occupation (1 = Employed full time, 2 = Employed part time, 3 = Unemployed looking for work, 4 = Unemployed not looking for work, 5 = Retired, 6 = Student, 7 = Disabled), Education (1 = Less than high school, 2 = High school graduate, 3 = MBO, 4 = Bachelor, 5 = Master, 6 = Doctorate). It should be noted that if no participants have given a certain answer possibility, that answer possibility is not included in this overview.

In *Table 9*, an overview of the descriptive statistics of the control variables within the filled in surveys of the pretest is given. From that table, one is able to notice that the survey was filled in by almost as much males as females, namely 26 males and 25 females. Most participants belonged to the age group 26-35 (N = 30), where all other participants belonged to the age group 16-25 (N = 21). Additionally, 76.5 percent of all participants (N = 39) denoted that they are full time employees, while 15.7 percent (N = 8) were students. In general, the group of participants could be considered as highly educated as 88.2 percent (N = 45) have obtained either a Bachelor or Master degree. The impact of having a sample set with a specific age group (all between 16 and 35), occupation (full time employees) and obtained degree (Bachelor or Master) will be discussed in the *Limitations* section of the chapter *Conclusion and Discussion* of this study.

4.2   Reviewer expertise

The objective of the pretest survey was to figure out which of the considered reviewers profiles in Appendix A was seen as the best option to represent either a reviewer profile with a high reputation or a reviewer profile with the low reputation. These certain two reviewer profiles are included in the survey of the main experiment to increase the likelihood that participants of that survey will also see a reviewer profile as either having a high or low reputation. In *Table 10*, an overview is given of the descriptive statistics of the considered reviewer profiles.

Reviewer	Min.	Max.	Mean	Standard Deviation
H1: 'Travel expert', profile picture and verified account	3	5	4.94	0.311
H2: 'Travel expert' and profile picture	3	5	4.00	0.447
H3: 'Travel expert'	2	5	3.39	0.635
L1: 'Travel amateur'	1	3	1.20	0.491
L2: 'Travel amateur' and profile picture	1	4	2.16	0.505
L3: 'Travel amateur', profile picture and verified account	2	5	2.76	0.651

Table 10: An overview of the descriptive statistics of the three different reviewers with a high reputation (H) and three different reviewers with a low reputation (L) within the pretest survey. Please note that the '#' in the first column denotes which specific reviewer is considered. Additionally, please note that reviewer expertise is measured on a five point Likert scale anchored at 1 = very low and 5 = very high.

From the table above, one is able to notice that reviewer profile H1, which has a 'travel expert' denotation, profile picture of a human included and a verified account, has the highest overall mean (4.94), while the reviewer profile L1, which only has a 'Travel amateur' denotation has the lowest overall mean (1.20). Thereby, one could conclude that these two profiles are best suited to represent a reviewer with either a high or low reputation in the survey of the main experiment. However, besides testing for significance, such conclusion

would be wrong, as participants were required to answer this question by a five point Likert scale anchored at 1 = very low and 5 = very high in order to create a full Likert scale, as noted in *Section 3.3*. This implies that the answer possibilities 'high' and 'very high', and 'low' and 'very low' should be considered as equals, because the reviewer expertise in this study is constructed as having a high or low reputation only, whereas no difference is made in how high (somewhat high, high or very high) or low that reputation is. Therefore, the reviewer profiles answer possibilities are recoded such that 'high' and 'very high' are grouped into one answer possibility 'high', and 'low' and 'very low' are grouped into one answer possibility 'low'. By such construction, one is able to compare the means of the different reviewer profiles to each other and decide which high and low reviewer profile should be included in the survey of the main experiment.

Reviewer expertise	Min.	Max.	Mean	Standard Deviation	t	Sig.
					Test value = 3	
H1: 'Travel expert', profile picture and verified account	2	3	2.98	0.140	-1.000	0.322
H2: 'Travel expert' and profile picture	2	3	2.90	0.300	-0.331	0.024*
H3: 'Travel expert'	1	3	2.35	0.559	-8.260	0.000*
					Test value = 1	
L1: 'Travel amateur'	1	2	1.03	0.196	1.429	0.159
L2: 'Travel amateur' and profile picture	1	3	1.17	0.477	2.638	0.011*
L3: 'Travel amateur', profile picture and verified account	1	3	1.74	0.594	8.497	0.000*

Table 11: An overview of the descriptive statistics and test output of the three different and transformed reviewers with a high reputation (H) and three different and transformed reviewers with a low reputation (L) within the pretest survey. Please note that the '#' in the first column denotes which specific reviewer is considered. Additionally, please note that all variables are measured on a five point Likert scale anchored at 1 = very low and 5 = very high and were transformed such that 'very low' and 'low', and 'high' and 'very high' were combined. It should be noted that (*) indicates a statistically significant result (p < 0.05, t = 1.96).

From *Table 11*, it can be seen that the means of all reviewer profiles are now closer to each other in comparison to the means in *Table 10*. Such finding can be explained as the answer possibilities '1 = very low' and '5 = very high' were transformed to either '2 = low' and '4 = high'. Reviewer profile H1 and L1 still have the highest and lowest overall mean, but it is necessary to test for significance. To explain such necessity, if the results of this question would show that one reviewer profile constructed as to have a high reputation would have a mean of 3, this would imply that only this one reviewer profile was denoted by all participants as having a high reputation. However, as is noticeable from *Table 11*, none of the means of the three considered reviewer profiles with a

high reputation has a value of exactly 3 (as similar for having a low reputation and considering the value of 1), implying that no reviewer profile was always seen by participants as having a high or low reputation.

To see if there exists a significant difference exists in how participants rated the level of expertise of the reviewers, the different means of the reviewer profiles in *Table 11* are compared to each other. Each mean of the reviewer profiles with a high reputation (H1: 2.98, H2: 2.90, H3: 2.35) is compared to each mean of the reviewer profiles with a low reputation (L1: 1.03, L2: 1.17, L3: 1.74). For example, the mean of the reviewer profile H1 is compared to the mean of reviewer profiles L1, L2 and L3. When comparing all these means, a significant difference is found for all comparisons ($p < 0.05$). This implies that the reviewer expertise of reviewer profiles with a high reputation was rated significantly different by participants in comparison to all reviewer profiles with a low reputation. It should be noted that when the means of *Table 10*, which were not transformed, are compared in a similar way, identical results were found.

An one sample *t*-test is used to test whether the obtained sample means individually significantly differ from their respective hypothetical mean of 3 (H_0: μ_{High} = 3) or 1 (H_0: μ_{Low} = 1). From the results in sixth and seventh column in *Table 11*, it can be noticed that the mean of one particular reviewer profile constructed as having a high reputation is not significantly different from 3 ($p < 0.05$, $t = 1.96$), namely the profile with a 'Travel expert' denotation, a profile picture of a human and a verified account. Similarly, only the reviewer profile with a 'Travel amateur' denotation does not has a significantly different mean from 1. Therefore, the latter profile is included in the survey of the main experiment as a reviewer profile with a low reputation, while the former profile is included as a reviewer profile with a high reputation.

4.3   Degree of negativity

Beforehand, three different less negative and three different very negative reviews were constructed, where each less negative review has a very negative counterpart to maintain consistency and continuity. Such combination of a less negative and very negative review only differs in the degree of negativity, but considers the same topics. The outcome of the pretest should show which combination of a very and less negative review should be included in the survey of the main experiment. This combination is included to increase the likelihood that participants of that survey will see a review as either very or less negative. In *Table 12*, an overview is given of the descriptive statistics of the three different less negative reviews as their three very negative counterparts, which are denoted by combination 1, 2 and 3.

From *Table 12*, it can be noticed that not one of the reviews is seen as 'positive' or 'very positive' by the participants. Some participants did see the reviews as 'neither negative or positive', but not one participant thought any review was (very) positively written. In addition, the three very negative reviews were always seen as either 'very negative' or 'negative'. Such finding is also noticeable when comparing the means of the less negative versus very negative reviews, where the means of the latter reviews are quite lower than the means of the less negative reviews. In order to test which combination of a less and very negative review should be

included in the survey of the main experiment, the means of the different reviews are compared and tested for significant difference.

Degree of negativity	Min.	Max.	Mean	Standard Deviation	*t*	Sig.
						Test value = 2
Less negative - combination 1	1	2	1.98	0.140	-1.000	0.322
Less negative - combination 2	1	3	1.76	0.513	-3.273	0.002*
Less negative - combination 3	1	3	2.00	0.400	0.000	1.000
						Test value = 1
Very negative - combination 1	1	2	1.06	0.238	1.768	0.083
Very negative - combination 2	1	2	1.02	0.140	1.000	0.322
Very negative - combination 3	1	2	1.08	0.272	2.063	0.044*

Table 12: An overview of the descriptive statistics and test output of the three different considered less negative reviews (LN) and three different considered very negative reviews (VN) within the pretest survey. Please note that the '#' in the first column denotes which specific review is considered. Additionally, please note that all variables are measured on a five point Likert scale anchored at 1 = very negative and 5 = very positive. It should be noted that (*) indicates a statistically significant result (p < 0.05, t = 1.96).

To see if participants rate the degree of negativity of less negative reviews significantly different from very negative reviews, the different means of the reviews in *Table 12* are compared to each other. Within each combination, the mean of the less negative review is compared to the mean of the very negative review. For example, within the first combination the mean of the less negative review (1.98) is compared to the mean of the very negative review (1.06). When comparing these means, a significant difference is found within all three combinations (p < 0.05). This implies that in all combinations, participants rated the degree of negativity of the less negative review significantly different from the very negative review.

An one sample *t*-test is used to test whether the obtained sample means individually significantly differ from their respective hypothetical mean. It was noted in *Section 4.2* that in a perfect world, all participants were able to denote one specific reviewer profile as either having a high or low reputation. Such reasoning is also used to test which combination of negative reviews should be included in the survey of the main experiment. Similar to testing for the reviewer profiles, a one sample t-test is used to test whether the obtained sample means significantly differ from their respective hypothetical mean. The hypothetical mean (or test value) for less negative reviews is 2 (H_0: μ_{Less} = 2), while the hypothetical mean for very negative reviews is 1 (H_0: μ_{Very} = 1). From the results in sixth and seventh column in *Table 12*, it can be observed that only one combination of a less and very negative review, namely the first one, does not have one review that has a statistically significant mean

from its respective hypothetical mean (p < 0.05). Therefore, the first combination of a less and very negative is included in the survey of the main experiment as a less and very negative review. However, it should be noted that the mean of the very negative review in this combination, when considering a 90% confidence level (p < 0.10, t = 1.645), is significantly different mean from its hypothetical mean. Considering the small sample size (N = 51), this would imply that if one or two participants had answered differently, when considering this confidence level, it might not be possible to denote a combination of negative reviews as suited to be used in the survey of the main experiment by only applying a one sample t-test. However, as a 95% confidence level is used throughout this study, the further impact of this obtained result will be discussed in the *Limitations* section of the chapter *Conclusion and Discussion.*

5   Results of the main experiment

5.1   Control variables

In *Table 13*, an overview of the descriptive statistics of the control variables within the filled in surveys of the main experiment is given. Participants were not only asked about their gender, age, current occupation and highest obtained degree, a question measuring their concern for the COVID-19 virus was also included. It has been decided that these variables are not considered in the analysis of the hypotheses, because than a very low amount of surveys would be compared to each other.

To give an example, only 12 female participants were shown experimental condition 1 (a trend seen over all 12 conditions as only 28.5% of the respondents of the main experiment were females), where at least 31 surveys (N > 30) would be needed to assume normally distributed data and conduct analysis. As this study's 12 conditions were only filled in by an average of around 40 respondents for each experimental condition, controlling for example age or gender by experimental condition would imply that a small amount of surveys (around 15-20 surveys depending on the control variable) would be compared to another small amount of surveys. Hence, the assumption of normally distributed data would be violated. Further implications of not controlling for variables as age and gender will be discussed in the *Limitations* section of the chapter *Conclusion and Discussion* of this study.

Variable	Min	Max	Mean (St. dev)
COVID-19	1	9	3.57 (1.287)
	Frequency (N = 494)	Percent	Cum. percent
Gender			
- Male	353	71.5	71.5
- Female	141	28.5	100.0
Age			
- 16/25	288	58.3	58.3
- 26/35	206	41.7	100.0
Occupation			
- Employed full time	170	34.4	34.4
- Employed part time	5	1.0	35.4
- Unemployed not looking for work	1	0.2	35.6
- Student	318	64.4	100.0
Degree			
- Less than high school	1	0.2	0.2
- High school graduate	5	1.0	1.2
- MBO	32	6.5	7.7

- Bachelor	206	41.7	49.4
- Master	249	50.4	99.8
- Doctorate	1	0.2	100.0

Table 13: An overview of the descriptive statistics of the control variables of the survey of the main experiment. Please note: COVID-19 (1 till 10, where a higher number indicates a higher concern for the virus), Gender (1 = Male, 2 = Female, 3 = Non-binary / third gender, 4 = Preferred not to say), Age (1 = Younger than 16 years old, 2 = 16 till 25, 3 = 26 till 35, 4 = 36 till 45, 5 = 46 till 55, 6 = 56 till 65, 7 = Older than 65 years old), Occupation (1 = Employed full time, 2 = Employed part time, 3 = Unemployed looking for work, 4 = Unemployed not looking for work, 5 = Retired, 6 = Student, 7 = Disabled), Education (1 = Less than high school, 2 = High school graduate, 3 = MBO, 4 = Bachelor, 5 = Master, 6 = Doctorate). It should be noted that if no participants have given a certain answer possibility, that answer possibility is not included in this overview.

In general, participants are not too concerned about the virus, as the mean average (on a scale of 1-10) is 3.57 (standard deviation = 1.287), where a higher number indicates a higher concern for the virus. Such finding seems to be in line with the latest research output about the rapidly declining concern virus in the Netherlands (RIVM, 2021). A total of 353 males and 141 females filled in the survey, which indicates that the distribution of male and female participants is quite different in comparison to the pretest. However, similar results are found when considering the age of the participants, as those also range from 16 till 35 years old. Additionally, 92.1% of the participants (N = 456) have obtained either a Bachelor, Master or Doctorate degree, which is similar to the highest obtained degree by participants of the pretest. Another noticeable difference is that, while the pretest survey was mostly filled in by full time employees, the survey of the main experiment is mostly filled in by students (64,2%). The impact of having a sample set, which noticeably differs in gender, age and occupation from the sample set of the pretest, will be discussed in the *Limitations* section of the chapter *Conclusion and Discussion* of this study.

5.2 Hypothesis testing

In order to test if the hypotheses hold, the following structure is considered. To start, it is tested if the interaction effect between degree of negativity and reviewer expertise moderates the relationship between offering compensation and the booking likelihood and hotel evaluation score of a reader. Such effect is tested for by three-way and two-way ANOVA tests. If such effect exists, the experimental conditions are selected based on the hypotheses. To give an example, when considering Hypothesis 1, the experimental conditions which offer full, partial and no compensation in response to very negative reviews written by a reviewer with a high reputation are selected. By use of a two-way ANOVA test (offering of compensation*interaction (degree of negativity*reviewer expertise)), it is tested if a significant difference exists between the means of the experimental conditions considered by each variable (hereafter referred to as 'between groups') by using pairwise comparisons.

Overall interaction effect

Before testing for the outcome of the hypotheses, the overall interaction effect is considered by use of a three-way ANOVA test. Such test is necessary, because it could be that the constructed interaction effect between reviewer expertise and degree of negativity does not have any influence on the relationship between offering compensation and the dependent variables. If so, this would imply that testing for the outcome of the hypotheses would be redundant, because all four hypotheses are constructed as such that the interaction effect between reviewer expertise and degree of negative does has a moderating effect on the relationship between offering compensation and booking likelihood and hotel evaluation score of a reader. *Table 14* gives an overview for the most important outcome of the three-way ANOVA test, which considers the three-way interaction effect between offering compensation, reviewer expertise and degree of negativity on the dependent variables booking likelihood, hotel evaluation score and recommendation. The full output of the three three-way ANOVA tests is included in Appendix B.

Three-way interaction	Booking likelihood	Hotel evaluation score	Recommendation
Compensation * Reviewer expertise * Degree of negativity	F = 4.333 Sig. = 0.070	F = 2.890 Sig. = 0.004*	F = 4.275 Sig. = 0.000*

Table 14: An overview of the output of the three-way ANOVA which considers the three-way effect between the offering compensation, reviewer expertise and degree of negativity on the dependent variables booking likelihood, hotel evaluation score and recommendation to friends and/or family. It should be noted that (*) indicates a statistically significant result (p < 0.05). Additionally, for the calculation of the F-test, the degrees of freedom of the interaction effect in all three cases is 8, while for the error: df = 459.

From the table above, it can be noticed that there is a statistically significant three-way interaction effect between compensation, reviewer expertise and degree of negativity on the dependent variables hotel evaluation score and recommendation ($p < 0.05$). In addition, when considering a 90% confidence level, such relationship also exists for the dependent variable booking likelihood. Thereby, it is shown that further testing for the outcome of the hypotheses is not redundant and is of value, although the three-way interaction effect on booking likelihood at a 90% confidence level. Additionally, it is found that the main effect of offering compensation on all considered dependent variables is significant, likewise to the main effects of both reviewer expertise and degree of negativity ($p < 0.05$). However, the two-way effect between offering compensation and degree of negativity is insignificant for hotel evaluation score and recommendation ($p > 0.05$), a finding in contrast to significant results obtained for booking likelihood ($p < 0.05$). Additionally, the two-way effect between reviewer expertise and degree of negativity is significant for hotel evaluation score and recommendation ($p < 0.05$), but not for booking likelihood ($p > 0.05$). To further analyse these results and test for the outcome of the hypotheses, a two-way ANOVA and pairwise comparison test are considered.

Two-way ANOVA test

To test for the outcome of the hypotheses and further look into the effects between offering compensation, degree of negativity and reviewer expertise on the dependent variables, a two-way ANOVA test is used. The constructed variable 'Interaction' has four levels based on the two levels degree of negativity and reviewer expertise. *Table 15* gives an overview for the most important outcome of the two-way ANOVA test, which considers the two-way interaction effect between offering compensation and interaction on the dependent variables booking likelihood, hotel evaluation score and recommendation.

Three-way interaction	Booking likelihood	Hotel evaluation score	Recommendation
Compensation * Interaction	F = 19.799	F = 14.443	F = 20.742
	Sig. = 0.000*	Sig. = 0.000*	Sig. = 0.000*

Table 15: An overview of the output of the two-way ANOVA which considers the two-way effect between the offering compensation and interaction on the dependent variables booking likelihood, hotel evaluation score and recommendation to friends and/or family. It should be noted that (*) indicates a statistically significant result ($p < 0.05$). Additionally, for the calculation of the F-test, the degrees of freedom of the interaction effect in all three cases is 6, while for the error: df = 481.

From *Table 15*, it can be seen that there is a statistically significant two-way interaction effect between compensation and interaction on all three dependent variables ($p < 0.05$). Hence, the constructed interaction between degree of negativity and reviewer expertise has influence on the relationship between offering compensation and the dependent variables. Additionally, it is found that the main effect of offering compensation on all considered dependent variables is significant, likewise to the main effect of interaction ($p < 0.05$). This further strengthens the believe that testing for the outcome of the hypotheses is of value.

Cronbach's Alpha

For each hypothesis, it is tested if the variables booking likelihood and recommendation to friends and/or family can be combined by use of Cronbach's Alpha. That measure shows how closely the answers of the participants to both questions are related. For example, it is assumed that participants who answer 'Strongly agree' to question 1 (Booking likelihood), will also answer 'Strongly agree' to question 3 (Recommendation). A higher Cronbach's Alpha score (>0.7) shows for such relation and implies that two (or more) variables can be combined.

By considering *Table 16*, it can be seen that for each hypothesis the variables booking likelihood and recommendation to friends and/or family can be combined into a new variable. The given answers are strongly related (as indicated by a Cronbach's Alpha > 0.9 for all hypotheses) and, therefore, the means of the two variables are combined and averaged over into a new variable, referred to as 'Booking likelihood'.

	Very negative review - High reputation (H1)	Less negative review - Low reputation (H2)	Very negative review - High reputation (H3)	Less negative review - Low reputation (H4)
Booking likelihood	3.30	4.16	3.89	4.40
Recommendation	3.32	4.14	3.90	4.40
Cronbach's Alpha	0.991	0.985	0.933	0.963
Combined mean	3.31	4.15	3.90	4.40

Table 16: An overview is given of the means, Cronbach's Alpha and combined mean for each hypothesis.

5.2.1 Hypothesis 1

Hypothesis 1: *Considering a very negative review written by a reviewer with a high reputation, the offering of full compensation leads to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while the offering of partial compensation leads to significantly higher booking likelihood and hotel evaluation score in comparison to not offering any compensation.*

To test for the outcome of this hypothesis, a two-way ANOVA test is used. From *Table 17*, the output of the two-way ANOVA test shows that there is a statistically significant simple effect between the offering of compensation on the booking likelihood (F = 121.977, $p < 0.05$) and hotel evaluation score (F = 96.775, $p < 0.05$) of a reader, when considering very negative reviews written by reviewers with a high reputation. Hence, in order to test which specific mean(s) differ(s), a pair wise comparison test is considered.

Variable	F	Sig.
Booking likelihood	121.977	0.000*
Evaluation score	96.775	0.000*

Table 17: The univariate test output of the two-way ANOVA test shows for the simple effect of offering compensation on the booking likelihood and hotel evaluation score of a reader, when considering very negative reviews written by reviewers with a high reputation. It should be noted that (*) indicates a statistically significant result ($p < 0.05$).

From *Table 18*, the pairwise comparison test reveals that there is a statistically significant difference in booking likelihood and hotel evaluation score, depending on which compensation is offered in response to a very negative review written by a reviewer with a high reputation. For both dependent variables, the offering of full compensation leads to a significantly higher booking likelihood and hotel evaluation score in comparison to offering partial ($p < 0.05$) or no compensation ($p < 0.05$). Additionally, the offering of partial compensation leads to a significantly higher result in comparison to not offering compensation at all ($p < 0.05$). Therefore, one can conclude that Hypothesis 1 holds, as the offering of full compensation does lead to a significantly higher booking

likelihood and evaluation score in comparison to partial compensation, while partial compensation does lead to a significantly higher booking likelihood and evaluation score in comparison to not offering any compensation.

Conditions	Booking likelihood			Hotel evaluation score		
	Means	Sig.	Outcome	Means	Sig.	Outcome
Full vs Partial	3.78 vs 2.95	0.000*	Full > Partial	6.61 vs 5.20	0.000*	Full > Partial
Full vs None	3.78 vs 1.51	0.000*	Full > No	6.61 vs 2.88	0.000*	Full > No
Partial vs None	2.95 vs 1.51	0.000*	Partial > No	5.20 vs 2.88	0.000*	Partial > No
		Overall	Full > Partial > No		Overall	Full > Partial > No

Table 18: An overview of the results of the pairwise comparison tests for Hypothesis 1. It should be noted that (*) indicates a statistically significant result (p < 0.05).

Similar results are found by observing the means of the variables response rate and response satisfaction segmented by level of compensation (*Table 30* of Appendix B). Responses which offer full compensation are significantly higher rated than ones which offer partial and no compensation (The mean of the response rate of offering full compensation (F) is 3.79, partial compensation (P) is 2.93 and no compensation (N) is 1.67). Likewise, when considering response satisfaction, full compensation offered is significantly higher rated than partial compensation , where the latter is significantly higher rated than no compensation (F: 3.78, P: 3.02, N: 1.66).

5.2.2 Hypothesis 2

Hypothesis 2*: Considering a less negative review written by a reviewer with a high reputation, the offering of full compensation does not lead to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while offering both levels of compensation does lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering of any compensation.*

To test for the outcome of this hypothesis, a two-way ANOVA test is used:

Variable	F	Sig.
Booking likelihood	107.343	0.000*
Evaluation score	76.962	0.000*

Table 19: The univariate test output of the two-way ANOVA test shows for the simple effect of offering compensation on the booking likelihood and hotel evaluation score of a reader, when considering less negative reviews written by reviewers with a high reputation. It should be noted that (*) indicates a statistically significant result (p < 0.05).

From *Table 19*, the output of the two-way ANOVA test shows that there is a statistically significant simple effect between the offering of compensation on the booking likelihood (F = 107.343, p < 0.05) and hotel evaluation score (F = 76.962, p < 0.05) of a reader, when considering less negative reviews written by reviewers with a high reputation. Hence, in order to test which specific mean(s) differ(s), a pair wise comparison test is considered.

	Booking likelihood			*Hotel evaluation score*		
Conditions	Means	Sig.	Outcome	Means	Sig.	Outcome
Full vs Partial	4.85 vs 4.62	0.086	Full = Partial	8.45 vs 7.93	0.046*	Full > Partial
Full vs None	4.85 vs 2.91	0.000*	Full > No	8.45 vs 5.37	0.000*	Full > No
Partial vs None	4.62 vs 2.91	0.000*	Partial > No	7.93 vs 5.37	0.000*	Partial > No
		Overall	Full = Partial > No		Overall	Full > Partial > No

Table 20: An overview of the results of the pairwise comparison tests for Hypothesis 2. It should be noted that (*) indicates a statistically significant result (p < 0.05).

When considering the output from the pairwise comparison test in *Table 20*, it can be seen that the booking likelihood only significantly differs between offering of compensation and not offering any compensation in response to a less negative review written by a reviewer with a high reputation. No significant difference in booking likelihood was found when comparing offering full and partial compensation to each other (p > 0.05). Offering of full and partial compensation lead to a significantly higher booking likelihood in comparison to offering no compensation at all (p < 0.05). However, when considering the output for the hotel evaluation score in *Table 20*, a significant difference between offering full and partial compensation is found (p < 0.05). Therefore, one is able to conclude that Hypothesis 2 only holds when considering the booking likelihood of participants, as the offering of full compensation in response to a less negative review written by a reviewer with a high reputation does lead to a significantly higher evaluation score in comparison to offering partial compensation.

When considering the results for the booking likelihood of a reader, similar results are found by observing the means of the variables response rate and response satisfaction (*Table 31* of Appendix B). Responses which offer full compensation are not significantly higher rated than ones which offer partial, but both are higher rated than no compensation offered (F: 4.77, P: 4.55, N: 2.85). Likewise, when considering response satisfaction, full compensation offered is not significantly higher rated than partial compensation, but both are significantly higher rated than no compensation (F: 4.80, P: 4.52, N: 3.12).

5.2.3 Hypothesis 3

Hypothesis 3: *Considering a very negative review written by a reviewer with a low reputation, the offering of full or partial compensation does not lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering any compensation.*

From *Table 21*, the output of the two-way ANOVA test shows that there is a statistically significant simple effect between the offering of compensation on the booking likelihood (F = 4.177, p < 0.05) and hotel evaluation score (F = 4.011, p < 0.05) of a reader, when considering very negative reviews written by reviewers with a low reputation. Hence, in order to test which specific mean(s) differ(s), a pair wise comparison test is considered.

Variable	F	Sig.
Booking likelihood	4.177	0.016*
Evaluation score	4.011	0.019*

Table 21: The univariate test output of the two-way ANOVA test shows for the simple effect of offering compensation on the booking likelihood and hotel evaluation score of a reader, when considering very negative reviews written by reviewers with a low reputation. It should be noted that (*) indicates a statistically significant result (p < 0.05).

From *Table 22*, the pairwise comparison test reveals that there is a statistically significant difference in booking likelihood and hotel evaluation score between the offering of full compensation and offering no compensation at all (p < 0.05). However, no significant difference in booking likelihood and evaluation score is found when comparing offering full compensation to partial compensation (p > 0.05). Additionally, the offering of partial compensation does not lead to significantly higher results in comparison to not offering compensation at all (p > 0.05). Thereby, Hypothesis 3 does not hold for both variables and, although the helpfulness of the reviews and trustworthiness of the reviewers of these three experimental conditions were lower rated than all others (as can be seen from *Tables 30 till 33* of Appendix B), the offering of full compensation is not regarded as overcompensation in comparison to not offering any compensation.

	Booking likelihood			*Hotel evaluation score*		
Conditions	Means	Sig.	Outcome	Means	Sig.	Outcome
Full vs Partial	4.10 vs 3.91	0.312	Full = Partial	6.98 vs 6.71	0.323	Full = Partial
Full vs None	4.10 vs 3.67	0.005*	Full > No	6.98 vs 6.22	0.005*	Full > No
Partial vs None	3.91 vs 3.67	0.066	Partial = No	6.71 vs 6.22	0.071	Partial = No
		Overall	-		Overall	-

Table 22: An overview of the results of the pairwise comparison tests for Hypothesis 3. It should be noted that (*) indicates a statistically significant result (p < 0.05).

Different results are found by observing the means of the variables response rate and response satisfaction (*Table 32* of Appendix B). Responses which offer full compensation are not significantly higher rated than ones which offer partial, but both are higher rated than no compensation offered (F: 4.77, P: 4.55, N: 3.12). Likewise,

when considering response satisfaction, full compensation offered is not significantly higher rated than partial compensation, but both are significantly higher rated than no compensation (F: 4.80, P: 4.52, N: 3.12).

5.2.4   Hypothesis 4

Hypothesis 4: *Considering a less negative review written by a reviewer with a low reputation, the offering of full compensation does not lead to a significantly higher booking likelihood and evaluation score in comparison to the offering of partial compensation, while offering both levels of compensation does lead to a significantly higher booking likelihood and hotel evaluation score in comparison to not offering of any compensation.*

From *Table 23*, the output of the two-way ANOVA test shows that there is a statistically significant simple effect between the offering of compensation on the booking likelihood (F = 22.047, p < 0.05) and hotel evaluation score (F = 17.513, p < 0.05) of a reader, when considering less negative reviews written by reviewers with a low reputation. Hence, in order to test which specific mean(s) differ(s), a pair wise comparison test is considered.

Variable	F	Sig.
Booking likelihood	22.047	0.000*
Evaluation score	17.513	0.000*

Table 23: The univariate test output of the two-way ANOVA test shows for the simple effect of offering compensation on the booking likelihood and hotel evaluation score of a reader, when considering less negative reviews written by reviewers with a low reputation. It should be noted that (*) indicates a statistically significant result (p < 0.05).

When considering the output from the pairwise comparison test in *Table 24*, it can be seen that the booking likelihood and evaluation score significantly differ between offering of compensation and not offering any compensation in response to a less negative review written by a reviewer with a low reputation. No significant difference in booking likelihood was found when comparing offering full and partial compensation to each other in both tests (p > 0.05). Offering of full and partial compensation lead to a significantly higher booking likelihood and hotel evaluation score in comparison to offering no compensation at all (p < 0.05). Therefore, one is able to conclude that Hypothesis 4 holds for both variables.

Similar results are found by observing the means of the variables response rate and response satisfaction segmented by level of compensation (*Table 33* of Appendix B). Responses which offer full compensation are not significantly higher rated than ones which offer partial, but both are higher rated than no compensation offered (F: 4.73, P: 4.66, N: 3.01). Likewise, when considering response satisfaction, full compensation offered is not significantly higher rated than partial compensation, but both are significantly higher rated than no compensation (F: 4.76, P: 4.46, N: 3.72).

	Booking likelihood			Hotel evaluation score		
Conditions	Means	Sig.	Outcome	Means	Sig.	Outcome
Full vs Partial	4.76 vs 4.58	0.526	Full = Partial	8.27 vs 8.10	0.132	Full = Partial
Full vs None	4.76 vs 3.83	0.000*	Full > No	8.27 vs 6.80	0.000*	Full > No
Partial vs None	4.58 vs 3.83	0.000*	Partial > No	8.10 vs 6.80	0.000*	Partial > No
		Overall	Full = Partial > No		Overall	Full = Partial > No

Table 24: An overview of the results of the pairwise comparison tests for Hypothesis 4. It should be noted that (*) indicates a statistically significant result (p < 0.05).

5.3 Manipulation check

The survey of the main experiment included two questions which were also asked in the survey of the pretest, namely about the level of expertise of the reviewer and the degree of negativity of the review. The latter was included as part of a manipulation check. The purpose of this check is to check if participants have actually read the review and response in the experimental condition. Surveys which failed to identify which compensation was offered by the hotel were already removed as part of the manipulation check. Moreover, it is tested if participants were able to successfully recognize if a less or very negative review and a reviewer with a low or high reputation was considered. Similar to testing of the degree of negativity and reviewer expertise in the pretest, a one sample t-test is used to test whether the obtained sample means significantly differ from their respective hypothetical mean. The hypothetical mean (or test value) for less negative reviews is 2 (H_0: μ_{Less} = 2), while the hypothetical mean for very negative reviews is 1 (H_0: μ_{Very} = 1), while the hypothetical mean for a reviewer with a low reputation is 1 (H_0: μ_{Low} = 1) and 3 for a reviewer with a high reputation (H_0: μ_{High} = 3).

From the results in sixth and seventh column in *Table 25*, it can be observed that, for both levels of negativity within the reviews, the observed means don't significantly differ from the hypothetical mean (p > 0.05). However, for both levels of reviewer expertise, the observed means do significantly differ from the hypothetical mean (p < 0.05). Therefore, one could argue that the manipulation check was only partially successful. Additionally, likewise to the analysis of the degree of negativity in the pretest, both means of the levels of degree of negativity would significantly differ when considering a 90% confidence level (p < 0.1). Although the sample size (N = 494) is larger than the sample size of the pretest (N = 51), this would imply that if some participants had answered differently, it might not have been possible to even conclude that the manipulation check partially successful. However, as a 95% confidence level is used throughout this study, the further impact of this obtained result will be discussed in the *Limitations* section of the chapter *Conclusion and Discussion.*

Variable	Min.	Max.	Mean	Standard Deviation	*t*	Sig.
Degree of negativity						*Test value = 2*
Less negative (N = 249)	1	3	1.96	0.302	-1.886	0.060
						Test value = 1
Very negative (N = 245)	1	3	1.02	0.168	1.900	0.059
Reviewer expertise						*Test value = 1*
Low reputation (N = 245)	1	2	1.34	0.688	7.814	0.000*
						Test value = 3
High reputation (N = 249)	2	3	2.87	0.461	-4.540	0.000*

Table 25: An overview of the descriptive statistics and test output of the less negative and very negative reviews and reviewers with a low or high reputation within the survey of the main experiment. Please note all variables were measured on a five point Likert scale anchored at 1 = very negative / very low and 5 = very positive / very high, while reviewer expertise was transformed such that 'very low' and 'low', and 'high' and 'very high' were combined. It should be noted that (*) indicates a statistically significant result (p < 0.05, t = 1,96).

To be sure that the manipulation check was successful and the results of this thesis are not invalid, a one-way ANOVA test is also used to test if participants were able to successfully distinct less negative reviews from very negative reviews and reviewers with a low reputation from reviewers with a high reputation. When comparing the means of less negative reviews and very negative (1.96 versus 1.02), the output of the one-way ANOVA test shows that a statistically significant difference between groups (F = 1820.949, p < 0.05) exists. Additionally, when comparing the transformed means of reviewers with a low reputation and reviewers with a high reputation (1.34 vs 2.87), the output of the one-way ANOVA test also shows that a statistically significant difference between groups (F = 837.118, p < 0.05) exists. Hence, it can be concluded that participants could successfully distinct between less negative and very negative reviews, and reviewers with a low reputation and high reputation.

6   Conclusion and Discussion

6.1   Contribution and implications

The objective of this study was to look into when and how hotels should offer compensation in response to a negative online review. Literature has produced mixed results on the impact of offering compensation in response to customer complaints in general, as negative online hotel reviews. This study focused on the relationship between the offering of compensation by a hotel in response to a negative online review, and the booking likelihood and hotel evaluation score of a reader, where the interaction effect between the review's degree of negativity and the reviewer's expertise is considered as moderator. Four hypotheses were constructed in order to test for the outcome of this study.

By results of the survey of the main experiment, it was shown that a three-way interaction effect between offering compensation, reviewer expertise and degree of negativity on the dependent variables hotel evaluation score, recommendation and booking likelihood (although at a 90% confidence level) exists. Additionally, it was shown that a two-way interaction effect between offering compensation and interaction (degree of negativity*reviewer expertise) on all three dependent variables also exists. However, it is found that not all constructed hypotheses are supported. Based on the results of this study's experimental design, hotel managers should offer compensation according to the following conditions if they want to achieve a higher booking likelihood and evaluation score of a reader and limit the impact of a negative online review. *Table 26* gives an overview of the final results of this study:

	High reputation	Low reputation
Very negative	H1: Full > Partial > No Result: F > P > N	H3: Full = Partial = No Result: F = P, P = N, F > N
Less negative	H2: Full = Partial > No Result: F = P > N (BL) / F > P > N (E)	H4: Full = Partial > No Result: F = P > N

Table 26: An overview of this study's results is given. It should be noted that 'Full', 'Partial' or 'No' corresponds to the level of offered compensation in a hotel response. Additionally, the indicator '>' denotes an outperformance in higher booking likelihood and hotel evaluation score by a certain level offered compensation in comparison to another level of compensation, while '=' denotes an equal performance in comparison to a different level of offered compensation. If the outcome of this study matches the constructed hypothesis, the result is given in green. If not, the result is given in red. It should be noted that (BL) applies to results considering the dependent variable booking likelihood, while (E) applies to results considering the dependent variable hotel evaluation score.

Given *Hypothesis 1*, it was assumed that a very negative review written by a reviewer with a high reputation would have a severe negative impact on the booking likelihood of a reader, while the review reputation of the reviewer would ensure that the review was seen as more helpful than a less negative review. Thereby, it was hypothesized that the offering of full compensation would lead to a significantly higher result than offering partial or no compensation, while partial compensation was expected to lead to higher results than offering no compensation. As can be seen in *Table 25*, this hypothesis was fully supported by the outcome of this study. Hotels would be best of offering full compensation in response to very negative reviews written by reviewers with a high reputation. When offered full compensation, the booking likelihood of a reader is close to 'high' and the hotel evaluation score is positive (mean = 6.61 (where all hotel evaluation scores were measured on a 10-point scale)). However, cash strapped hotels could also opt to offer partial compensation as, although significantly lower than offering full compensation, the offering of partial compensation in response to a very negative review written by a reviewer with a high reputation would lead to significantly higher results in booking likelihood and hotel evaluation score in comparison to not offering compensation at all. Yet, it should be noted that if partial compensation is offered, the booking likelihood and of a reader would only be close to 'neutral' (mean = 2.95) and the hotel evaluation score would still be negative (mean = 5.20), which raises questions if offering partial compensation in this setting should be considered at all given its limited impact. Moreover, not offering any compensation to very negative reviews written by a reviewer with a high reputation is absolutely not recommendable, as this study shows that the booking likelihood of a reader (mean = 1.51) and hotel evaluation score (mean = 2.88) would be very low.

Given *Hypothesis 2*, it was assumed that a less negative review written by a reviewer with a high reputation would have a less severe negative impact on the booking likelihood of a reader than a very negative review written by the same reviewer. This implied that it was assumed that offering full compensation would be overcompensating, due to the magnitude of the customer complaint (a less negative review). Thereby, it was hypothesized that offering full and partial compensation would lead to equally statistically significant higher results in comparison to not offering any compensation when considering a less negative review written by a reviewer with a high reputation. As can be seen in *Table 25*, this hypothesis was partially supported by the outcome of this study. If a hotel would find increasing the booking likelihood of a reader more important, hotels would be best of by offering compensation, as the booking likelihood of a reader of both levels of compensation is significantly higher than offering no compensation. However, as no significant difference was found between the means of both levels of compensation (means: 4.85 vs 4.62), hotels are better of offering partial compensation, as offering full compensation in this setting is overcompensating. In addition, hotels which want to focus on increasing their hotel evaluation score could also offer partial compensation (mean = 7.93), although that score is significantly lower than the hotel evaluation score of a reader when offered full compensation (mean = 8.45). Cash strapped hotels could consider offering no compensation to allocate financial resources to offering compensation to other negative reviews, as the booking likelihood and of a reader would be close to 'neutral' (mean = 2.91) and the hotel evaluation score would be slightly above '5' (mean = 5.37).

Given *Hypothesis 3*, it was assumed that readers were more likely to ignore very negative review written by reviewers with a low reputation and, therefore, the negative impact of these reviews on the booking likelihood of a reader would be very limited. Thereby, it was hypothesized that it doesn't matter which level of compensation was offered, the booking likelihood and hotel evaluation score of a reader would not be significantly different. As can be seen in *Table 25*, this hypothesis was not supported by the outcome of this study. Such finding can be explained as it is suggested by Yin et al. (2020) that readers ignore these kinds of reviews only if they have access to other reviews (as less negative or positive ones) about the same hotel, but in this study readers only had access to one certain review. Moreover, very mixed results were found in this setting, depending on which variable (booking likelihood, hotel evaluation score, response rate or response satisfaction) was considered. When considering the booking likelihood and hotel evaluation score of a reader, hotels would be best of offering full compensation in response to a very negative review written by a reviewer with a low reputation. Yet, when considering the response rate and response satisfaction hotels would be best of offering partial compensation, as no significant difference in mean scores of these variables was found in comparison to the offering of full compensation. Moreover, while this study is not able give a definite verdict on the offering of full or partial compensation in response to a very negative review written by reviewer with a low reputation, cash strapped companies can still consider offering no compensation at all. Although significantly lower than both offering full or partial compensation, the booking likelihood of a reader would be close to 'high' (mean = 3.67) and hotel evaluation score of a reader would be acceptable (mean = 6.21).

Given *Hypothesis 4*, it was assumed that less negative reviews written by a reviewer with a low reputation would be more likely to be seen as a true reflection of the hotel's performance and would have a larger negative impact on the booking likelihood of a reader than very negative reviews written by the same reviewer. Thereby, it was hypothesized that the offering of full and partial will lead to equally significant higher results in comparison to offering no compensation at all. As can be seen in *Table 25*, this hypothesis was fully supported by the outcome of this study. In this setting, hotels would be best of offering partial compensation in response, as offering full compensation does not lead to significant higher results. However and likewise to the previous hypothesis, cash strapped hotels could also consider not offering any compensation in response, because he booking likelihood of a reader would be close to 'high' (mean = 3.83) and hotel evaluation score of a reader would be very acceptable (mean = 6.80).

In general, hotels are advised to offer full compensation in response to a very negative review regardless of reviewer expertise. However, in real life, it is very probable that hotels do not have the financial resources to offer such compensation (Piehler et al., 2019) , which undermines the applicability of this advice. Therefore, based on the results of this study, cash strapped hotel are advised to at least offer partial compensation in response to a very negative review written by a reviewer with a high reputation.

6.2   Limitations

Although this research was conducted with careful considerations, there are some limitations to it. The most important limitation of this study is the sample set of participants of both the pretest and main experiment, which primarily includes youngsters (aged 16 till 35) and are highly educated (have a Bachelor degree or higher). Additionally, the sample set of the pretest differs from the sample set of the main experiment by occupation, as most participants in the pretest were students and equally divided over gender, while most participants in the main experiment are male full time employees. Therefore, the outcome of this study is only representative of these subgroups of the population and not the entire population. However, it should be noted that does not imply that the results obtained in this study are invalid. Both the pretest and the main experiment included a manipulation check, in order to test if the experimental design was successfully manipulated. In both cases, the manipulation check was successful, which implies that the results of this study can be considered as valid, regardless of differences in the sample set.

That participants are youngsters and are highly educated is caused by how data was collected. Both the pretest and main experiment were distributed over friends of the writer, which heavily influences which participants have filled in these surveys. Thereby, another limitation of this study is the limited scope of this study. While already quite complex by considering 12 experimental conditions, when this study's experimental design is compared to other experimental designs considered as sources for this study, this study's design mostly lacks in size of the sample sets (which is why control variables were excluded from the analysis of the hypotheses) and limited amount of included explanatory variables in the conceptual framework. To be more specific, the number of variables considered in the conceptual framework was limited, as this study is only able to obtain some hundreds of participants (and for each experimental condition $N > 30$ is required). To give an example of an additional explanatory variable, the type of hotel may influence the relationship between the offering of compensation and booking likelihood. It is argued by Li et al. (2017) and De Pelsmacker et al. (2018) that this relationship is potentially moderated by characteristics of a hotel, as the hotel's star rating and size, due to different customer expectations when considering luxury products as hotel rooms. Other variables which may influence this relationship are the characteristics of a review, as the number of considered reviews and the amount of times a review is marked as 'helpful' (Sparks et al., 2016). However, as the number of variables grows, the number of needed participants grows accordingly, which is why this study's conceptual framework was limited in the amount of variables considered.

6.3   Further research

To fully understand how the booking likelihood of consumers is influenced by the offering of compensation, further research would need to simulate a real word setting to perfection and considering more explanatory variables as for example type of hotel. As this study is limited in scale, it is suggested to construct a research design which perfectly mimics a real world booking website (as TripAdvisor) and ask participants to book a hotel. By doing qualitative research, every decision of the participant, from researching to booking a hotel, can be

followed by use of advanced techniques as eye tracking. Additionally, more explanatory variables can be considered, as booking websites consider for example all kinds of hotels, reviewers, reviews and responses by hotels. Such analysis can also answer to another limitation of this study: do participants actually read every review and response when multiple reviews and responses are available? And how does this impact their booking likelihood with a hotel? Finding an answer to these questions would be of utmost value to hotels and companies in general.

References

Ady, M., Quadri-Felitti, D. et al. (2015), 'Consumer research identifies how to present travel review content for more bookings', *Retrieved form http://webcache. googleusercontent. com/search* .

Akehurst, G. (2009), 'User generated content: the use of blogs for tourism organisations and tourism consumers',
Service business 3(1), 51.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D., Glover, G., Gabrieli, J. D. and Sobel,
N. (2003), 'Dissociated neural representations of intensity and valence in human olfaction', *Nature neuroscience* 6(2), 196–202.

Banerjee, S., Bhattacharyya, S. and Bose, I. (2017), 'Whose online reviews to trust? understanding reviewer trustworthiness and its impact on business', *Decision Support Systems* 96, 17–26.

Bowen, J., Baloglu, S., Lee, H. and Blum, S. C. (2015), 'How hotel responses to online reviews differ by hotel rating: an exploratory study', *Worldwide Hospitality and Tourism Themes* .

Browning, V., So, K. K. F. and Sparks, B. (2013), 'The influence of online reviews on consumers' attributions of service quality and control for service standards in hotels', *Journal of Travel & Tourism Marketing* 30(1-2), 23–40.

Cacioppo, J. T., Petty, R. E., Losch, M. E. and Kim, H. S. (1986), 'Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions.', *Journal of personality and social psychology* 50(2), 260.

Casado-Díaz, A. B., Andreu, L., Beckmann, S. C. and 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.

Chan, I. C. C., Lam, L. W., Chow, C. W., Fong, L. H. N. and Law, R. (2017), 'The effect of online reviews on hotel booking intention: The role of reader-reviewer similarity', *International Journal of Hospitality Management* 66, 54–65.

Chan, N. L. and Guillet, B. D. (2011), 'Investigation of social media marketing: how does the hotel industry in hong kong perform in marketing on social media websites?', *Journal of Travel & Tourism Marketing* 28(4), 345–368.

Chatterjee, P. (2001), 'Online reviews: do consumers use them?', *Association for Consumer Research* .

Cheung, C. M. and Thadani, D. R. (2012), 'The impact of electronic word-of-mouth communication: A literature analysis and integrative model', *Decision support systems* 54(1), 461–470.

Cheung, C. M.-Y., Sia, C.-L. and Kuan, K. K. (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(8), 2.

Chevalier, J. A. and Mayzlin, D. (2006), 'The effect of word of mouth on sales: Online book reviews', *Journal of marketing research* 43(3), 345–354.

Childers, T. L., Carr, C. L., Peck, J. and Carson, S. (2001), 'Hedonic and utilitarian motivations for online retail shopping behavior', *Journal of retailing* 77(4), 511–535.

Chua, A. Y. and 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(2), 354–362.

Conlon, D. E. and Murray, N. M. (1996), 'Customer perceptions of corporate responses to product complaints: The role of explanations', *Academy of management journal* 39(4), 1040–1056.

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

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

Dixon, T. and Marston, A. (2002), 'Uk retail real estate and the effects of online shopping', *Journal of Urban Technology* 9(3), 19–47.

Duan, W., Gu, B. and Whinston, A. B. (2008), 'Do online reviews matter?—an empirical investigation of panel data', *Decision support systems* 45(4), 1007–1016.

Ekiz, E., Khoo-Lattimore, C. and Memarzadeh, F. (2012), 'Air the anger: investigating online complaints on luxury hotels', *Journal of Hospitality and Tourism Technology* .

El-Manstrly, D. and Wirtz, J. (2020), 'Proactive handling of flight overbooking: How to reduce negative ewom and the costs of bumping customers', *Journal of Service Research* 1, 20.

Estelami, H. (2000), 'Competitive and procedural determinants of delight and disappointment in consumer complaint outcomes', *Journal of service research* 2(3), 285–300.

Farrell, M. J. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society: Series A (General)* 120(3), 253–281.

Feldman, S. (1966), Motivational aspects of attitudinal elements and their place in cognitive interaction, *in* 'Cognitive consistency', Elsevier, pp. 75–108.

Fick, G. R. and Ritchie, J. (1991), 'Measuring service quality in the travel and tourism industry', *Journal of travel research* 30(2), 2–9.

Flanagin, A. J. and Metzger, M. J. (2013), 'Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics', *Computers in Human Behavior* 29(4), 1626–1634.

Fornell, C. and Wernerfelt, B. (1987), 'Defensive marketing strategy by customer complaint management: a theoretical analysis', *Journal of Marketing research* 24(4), 337–346.

Garrett, D. E. (1999), 'The effectiveness of compensation given to complaining consumers: is more better?', *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior* 12.

Gelbrich, K. and Roschk, H. (2011), 'Do complainants appreciate overcompensation? a meta-analysis on the effect of simple compensation vs. overcompensation on post-complaint satisfaction', *Marketing Letters* 22(1), 31–47.

Goodwin, C. and Ross, I. (1989), 'Salient dimensions of perceived fairness in resolution of service complaints', *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior* 2(14), 87–92.

Gruen, T. W., Osmonbekov, T. and Czaplewski, A. J. (2006), 'ewom: The impact of customer-to-customer online know-how exchange on customer value and loyalty', *Journal of Business research* 59(4), 449–456.

Gvili, Y. and Levy, S. (2018), 'Consumer engagement with ewom on social media: the role of social capital', *Online Information Review* .

Hayes, D. and Lawless, J. L. (2018), 'The decline of local news and its effects: New evidence from longitudinal data', *The Journal of Politics* 80(1), 332–336.

Heiman, R. L. (1997), 'The impact of the restaurant critic', *Hospitality Review* 15(2), 10.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G. and Gremler, D. D. (2004), 'Electronic word-of-mouth via consumeropinion platforms: what motivates consumers to articulate themselves on the internet?', *Journal of interactive marketing* 18(1), 38–52.

Hicks, A., Comp, S., Horovitz, J., Hovarter, M., Miki, M. and Bevan, J. L. (2012), 'Why people use yelp. com: An exploration of uses and gratifications', *Computers in Human Behavior* 28(6), 2274–2279.

Hoffman, K. D., Kelley, S. W. and Rotalsky, H. M. (1995), 'Tracking service failures and employee recovery efforts', *Journal of services marketing* .

Hu, N., Liu, L. and Zhang, J. J. (2008), 'Do online reviews affect product sales? the role of reviewer characteristics and temporal effects', *Information Technology and management* 9(3), 201–214.

Hua, W., Chan, A. and Mao, Z. (2009), 'Critical success factors and customer expectation in budget hotel segment—a case study of china', *Journal of Quality Assurance in Hospitality & Tourism* 10(1), 59–74.

Hudson, S. and Thal, K. (2013), 'The impact of social media on the consumer decision process: Implications for tourism marketing', *Journal of Travel & Tourism Marketing* 30(1-2), 156–160.

Hussain, S., Ahmed, W., Jafar, R. M. S., Rabnawaz, A. and Jianzhou, Y. (2017), 'ewom source credibility, perceived risk and food product customer's information adoption', *Computers in Human Behavior* 66, 96–102.

Istanbulluoglu, D. (2017), 'Complaint handling on social media: The impact of multiple response times on consumer satisfaction', *Computers in Human Behavior* 74, 72–82.

Janiszewski, C. and Lichtenstein, D. R. (1999), 'A range theory account of price perception', *Journal of Consumer Research* 25(4), 353–368.

Jensen, M. L., Averbeck, J. M., Zhang, Z. and Wright, K. B. (2013), 'Credibility of anonymous online product reviews: A language expectancy perspective', *Journal of Management Information Systems* 30(1), 293–324.

Kim, T. T., Kim, W. G. and Kim, H.-B. (2009), 'The effects of perceived justice on recovery satisfaction, trust, word-of-mouth, and revisit intention in upscale hotels', *Tourism management* 30(1), 51–62.

Kobez, M. (2018), "restaurant reviews aren't what they used to be': digital disruption and the transformation of the role of the food critic', *Communication Research and Practice* 4(3), 261–276.

Kusumasondjaja, S., Shanka, T. and Marchegiani, C. (2012), 'Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence', *Journal of Vacation Marketing* 18(3), 185–195.

Kwok, L. and Xie, K. L. (2016), 'Factors contributing to the helpfulness of online hotel reviews', *International Journal of Contemporary Hospitality Management* .

Kwok, L., Xie, K. L. and Richards, T. (2017), 'Thematic framework of online review research', *International Journal of Contemporary Hospitality Management* .

Lee, M., Jeong, M. and Lee, J. (2017), 'Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website', *International Journal of Contemporary Hospitality Management* .

Lee, M. and Youn, S. (2009), 'Electronic word of mouth (ewom) how ewom platforms influence consumer product judgement', *International Journal of Advertising* 28(3), 473–499.

Leung, X. Y., Sun, J. and Bai, B. (2019), 'Thematic framework of social media research: state of the art', *Tourism Review* .

Levy, S. E., Duan, W. and Boo, S. (2013), 'An analysis of one-star online reviews and responses in the washington, dc, lodging market', *Cornell Hospitality Quarterly* 54(1), 49–63.

Lewis, R. C. (1983), *Consumers complain-what happens when business responds?*, Bloomington.

Li, C., Cui, G. and Peng, L. (2017), 'The signaling effect of management response in engaging customers: A study of the hotel industry', *Tourism Management* 62, 42–53.

Lindsey-Mullikin, J. and Borin, N. (2017), 'Why strategy is key for successful social media sales', *Business Horizons* 60(4), 473–482.

Liu, H., Jayawardhena, C., Dibb, S. and 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, Y., Song, Y., Sun, J., Sun, C., Liu, C. and Chen, X. (2020), 'Understanding the relationship between food experiential quality and customer dining satisfaction: A perspective on negative bias', *International Journal of Hospitality Management* 87, 102381.

Lo, A. S. and Yao, S. S. (2019), 'What makes hotel online reviews credible?', *International Journal of Contemporary Hospitality Management* .

Luca, M. (2016), 'Reviews, reputation, and revenue: The case of yelp. com', *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper* (12-016).

Mattila, A. S. and Mount, D. J. (2003), 'The impact of selected customer characteristics and response time on e-complaint satisfaction and return intent', *International journal of hospitality management* 22(2), 135–145.

Mattila, A. S. and Ro, H. (2008), 'Discrete negative emotions and customer dissatisfaction responses in a casual restaurant setting', *Journal of Hospitality & Tourism Research* 32(1), 89–107.

Mauri, A. G. and Minazzi, R. (2013), 'Web reviews influence on expectations and purchasing intentions of hotel potential customers', *International journal of hospitality management* 34, 99–107.

McGinnies, E. and Ward, C. D. (1980), 'Better liked than right: Trustworthiness and expertise as factors in credibility', *Personality and Social Psychology Bulletin* 6(3), 467–472.

Metzger, M. J., Flanagin, A. J. and Medders, R. B. (2010), 'Social and heuristic approaches to credibility evaluation online', *Journal of communication* 60(3), 413–439.

Min, H., Lim, Y. and 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.

Mizerski, R. W. (1982), 'An attribution explanation of the disproportionate influence of unfavorable information', *Journal of consumer research* 9(3), 301–310.

Mohsin, A., Rodrigues, H. and Brochado, A. (2019), 'Shine bright like a star: Hotel performance and guests' expectations based on star ratings', *International Journal of Hospitality Management* 83, 103–114.

Morris, S. V. (1988), 'How many lost customers have you won back today? an aggressive approach to complaint handling in the hotel industry', *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior* 1(1), 86–92.

Mount, D. J. and Mattila, A. (2000), 'The final opportunity: The effectiveness of a customer relations call center in recovering hotel guests', *Journal of Hospitality & Tourism Research* 24(4), 514–525.

Namkung, Y., Jang, S. S. and Choi, S. K. (2011), 'Customer complaints in restaurants: Do they differ by service stages and loyalty levels?', *International Journal of Hospitality Management* 30(3), 495–502.

Noone, B. M., & Lee, C. H. (2011). Hotel overbooking: The effect of overcompensation on customers' reactions to denied service. *Journal of Hospitality & Tourism Research*, *35*(3), 334-357.

Park, C. and Lee, T. M. (2009), 'Information direction, website reputation and ewom effect: A moderating role of product type', *Journal of Business research* 62(1), 61–67.

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

Pavitt, D. (1997), 'Retailing and the super high street: the future of the electronic home shopping industry', *International Journal of Retail & Distribution Management* .

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

Piehler, R., Schade, M., Hanisch, I. and Burmann, C. (2019), 'Reacting to negative online customer reviews', *Journal of Service Theory and Practice* .

Poelman, M. P., Gillebaart, M., Schlinkert, C., Dijkstra, S. C., Derksen, E., Mensink, F., Hermans, R. C., Aardening, P., de Ridder, D. and de Vet, E. (2020), 'Eating behavior and food purchases during the covid-19 lockdown: A cross-sectional study among adults in the Netherlands', *Appetite* 157, 105002.

RIVM (2021), 'Naleven gedragsregels', https://www.rivm.nl/gedragsonderzoek/maatregelen-welbevinden/naleven-gedragsregels

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

Schoenmüller, V., Netzer, O. and Stahl, F. (2019), 'The extreme distribution of online reviews: Prevalence, drivers and implications', *Columbia Business School Research Paper* 18(10).

Schouten, J. (2020), De fysieke winkel is dood, lang leve de fysieke winkel!, Master's thesis.

Schroeder, J. J. (1985), 'Restaurant critics respond: We're doing our job', *Cornell Hotel and Restaurant Administration Quarterly* 25(4), 56–63.

Schuckert, M., Liu, X. and Law, R. (2015), 'Hospitality and tourism online reviews: Recent trends and future directions', *Journal of Travel & Tourism Marketing* 32(5), 608–621.

Sen, S. and Lerman, D. (2007), 'Why are you telling me this? an examination into negative consumer reviews on the web', *Journal of interactive marketing* 21(4), 76–94.

Sheng, J., Amankwah-Amoah, J., Wang, X. and Khan, Z. (2019), 'Managerial responses to online reviews: A text analytics approach', *British Journal of Management* 30(2), 315–327.

Sotiriadis, M. D. (2017), 'Sharing tourism experiences in social media', *International Journal of Contemporary Hospitality Management* .

Sparks, B. A. and Bradley, G. L. (2014), 'A "triple a" typology of responding to negative consumer-generated online reviews', *Journal of Hospitality & Tourism Research* 41(6), 719–745.

Sparks, B. A. and Browning, V. (2011), 'The impact of online reviews on hotel booking intentions and perception of trust', *Tourism management* 32(6), 1310–1323.

Sparks, B. A., So, K. K. F. and 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.

Tax, S. S., Brown, S. W. and Chandrashekaran, M. (1998), 'Customer evaluations of service complaint experiences: implications for relationship marketing', *Journal of marketing* 62(2), 60–76.

Trappey, A. J., Lee, C.-H., Chen, W.-P. and Trappey, C. V. (2010), A framework of customer complaint handling system, *in* '2010 7th International Conference on Service Systems and Service Management', IEEE, pp. 1–6.

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

Voss, G. B., Parasuraman, A. and Grewal, D. (1998), 'The roles of price, performance, and expectations in determining satisfaction in service exchanges', *Journal of marketing* 62(4), 46–61.

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

Wolfinbarger, M. and Gilly, M. C. (2001), 'Shopping online for freedom, control, and fun', *California management review* 43(2), 34–55.

Wu, J., Wu, T. and Schlegelmilch, B. B. (2020), 'Seize the day: how online retailers should respond to positive reviews', *Journal of Interactive Marketing* 52, 52–60.

Xiang, Z. and Gretzel, U. (2010), 'Role of social media in online travel information search', *Tourism management* 31(2), 179–188.

Xie, H. J., Miao, L., Kuo, P.-J. and Lee, B.-Y. (2011), 'Consumers' responses to ambivalent online hotel reviews: The role of perceived source credibility and pre-decisional disposition', *International Journal of Hospitality Management* 30(1), 178–183.

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

Xu, Q. (2014), 'Should i trust him? the effects of reviewer profile characteristics on ewom credibility', *Computers in Human Behavior* 33, 136–144.

y Monsuw´e, T. P., Dellaert, B. G. and De Ruyter, K. (2004), 'What drives consumers to shop online? a literature review', *International journal of service industry management* .

Ye, Q., Law, R. and Gu, B. (2009), 'The impact of online user reviews on hotel room sales', *International Journal of Hospitality Management* 28(1), 180–182.

Yin, D., Bond, S. and Zhang, H. (2020), 'Anger in consumer reviews: Unhelpful but persuasive?', *MIS Quarterly, forthcoming*

Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews?. *Information Technology & Tourism*, *10*(4), 283-295.

Zeelenberg, M. and Pieters, R. (2004), 'Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services', *Journal of business Research* 57(4), 445–455.

Zhang, J. and Wedel, M. (2009), 'The effectiveness of customized promotions in online and offline stores', *Journal of marketing research* 46(2), 190–206.

Zhang, Z., Ye, Q., Law, R. and Li, Y. (2010), 'The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews', *International Journal of Hospitality Management* 29(4), 694–700.

Zhao, Y., Xu, X. and Wang, M. (2019), 'Predicting overall customer satisfaction: Big data evidence from hotel online textual reviews', *International Journal of Hospitality Management* 76, 111–121.

Appendix A

A: Reviewer profiles in pretest

B: Reviews in pretest

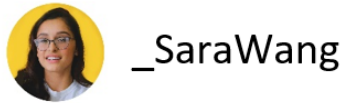**C: Control questions (pretest)**

D: Control questions (main experiment)

E: Hotel responses

A. Considered reviewer profiles in the pretest (order: High reputation profile #1, High reputation profile #2, High reputation profile #3, Low reputation profile #1, Low reputation profile #2, Low reputation profile #3):

_SaraWang ✓

Travel expert • 1986 contributions • 4637 helpful votes

_SaraWang

Travel expert • 1986 contributions • 4637 helpful votes

S _SaraWang

Travel expert • 1986 contributions • 4637 helpful votes

S _SaraWang

Travel amateur • 1 contribution • 0 helpful votes

_SaraWang

Travel amateur • 1 contribution • 0 helpful votes

_SaraWang ✓
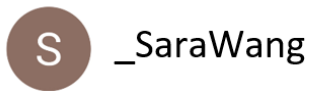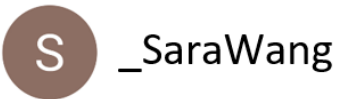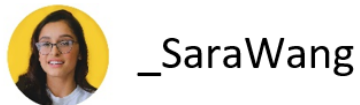
Travel amateur • 1 contribution • 0 helpful votes

B. *Considered reviews in the pretest (order: less negative 1, less negative 2, less negative 3, very negative 1, very negative 2, very negative 3):*

★ ★☆ ☆ ☆

Great location but uncomfortable stay

Stayed here for a few nights. Found the location to be close to the city centre and surrounding areas, which was very convenient. We paid for an iffy breakfast and lunch package. The breakfast was awful, and we opted on not using our breakfast credits after we went there once. In addition, the premade cocktails were pretty bad. As for the room, it was not cleaned beforehand and the bed was very uncomfortable. All in all, there are better hotels to consider.

★ ★☆ ☆ ☆

Could use some curtains

The building across the street is under construction and they have floodlights on. Wouldn't be an issue if there were blackout curtains, but there are only blinds, which do very little. Besides that the sheets and blankets are on the old side of the spectrum and the carpet is a bit grimey. Location is nice, staff is nice, but unless you like sleeping in a bit pricy and tiny box, maybe pass on this one?

★ ★☆ ☆ ☆

Some positive, mostly negative

The hotel is at a supreme location but very expensive in terms of money which we paid. The staff wasn't supportive and a bit rude. The rooms are not having much space to move with luggage and wash rooms are not attached. It is having a common toilet and it was not very comfortable. Drinking water cooler was not functioning. I was there to attend a conference with my wife. Booked online but later on realized that in this rate, we were having many options and get much better hotel.

★☆ ☆ ☆ ☆

Do not stay at this hotel!

Stayed here for a few nights. Found the location to be very far away from the city centre and surrounding areas. We paid for a horrible breakfast and lunch package. The breakfast was ABSOLUTELY disgusting, and we opted on not using our credits ever again after we went there once. In addition, the premade cocktails were unquestionably overdue. To make matters worse, our room was disgusting and the bed was very uncomfortable... All in all, an absolute waste of money!

★☆ ☆ ☆ ☆

Please don't go

The building across the street is under construction and they have floodlights on. Wouldn't be an issue if there were blackout curtains, but there are only blinds, which do ABSOLUTELY nothing. Besides that the sheets and blankets are from the middle ages and the carpet is disgusting. Location is shit, staff is worse and unless you like sleeping in a pricy radioactively bright torture box, you should definitely pass on this one!

★☆ ☆ ☆ ☆

Way better options available

The hotel is at an awful location and very expensive in terms of money which we paid. The staff was not supportive at all and VERY rude. The rooms are not having much space to move with luggage and wash rooms are not attached. It is having a common toilet and it was horrible. Drinking water cooler was not functioning. I was there to attend a conference with my wife. Booked online but later on realized that in this rate, we were having many options and get much better hotel!!!

C. *Control questions in the pretest:*

To which gender do you identify?

Answer possibilities: Male, Female, Non-binary / Third gender, Prefer not to say

How old are you?

Answer possibilities: Younger than 16 years old, 16-25, 26-35, 36-45, 46-55, 56-65, Older than 65 years old

What is your current occupation?

Answer possibilities: Employed full time, Employed part time, Unemployed looking for work, Unemployed not looking for work, Retired, Student, Disabled

What is the highest degree you obtained?

Answer possibilities: Less than high school, High school graduate, MBO, Bachelor, Master, Doctorate

D. *Control questions in the main experiment:*

On a scale of 1 till 10, how concerned are you about the COVID-19 virus? Please note that a higher number denotes a stronger concern for the virus.

Answer possibilities: 1-10

To which gender do you identify?

Answer possibilities: Male, Female, Non-binary / Third gender, Prefer not to say

How old are you?

Answer possibilities: Younger than 16 years old, 16-25, 26-35, 36-45, 46-55, 56-65, Older than 65 years old

What is your current occupation?

Answer possibilities: Employed full time, Employed part time, Unemployed looking for work, Unemployed not looking for work, Retired, Student, Disabled

What is the highest degree you obtained?

Answer possibilities: Less than high school, High school graduate, MBO, Bachelor, Master, Doctorate

E. Considered responses in the main experiment (order full, partial, no; where the difference in offered compensation is <u>highlighted</u>):

Dear Sara,

Thank you for taking the time to write a review of your recent stay at our hotel. We greatly appreciate your comments.

It was concerning to read about the issues you encountered during your recent stay and I would like to apologize for the inconvenience it caused. I have personally inspected your room and it seems that we dropped the ball in a few areas during your visit.

Unfortunately during your stay, the city council had ordered a partial lockdown due to the COVID-19 pandemic. As we informed you, this implied that we were working with less than half of our staff. I have met with our whole staff to ensure that any further inconvenience will be minimized in the future.

Although we try to create the best experience possible for our hotel guests during a global pandemic, we endeavour to meet the room requirements of all our guests at all times. Thereby, we would like to <u>fully compensate</u> you for your stay at our hotel and hope that you will visit us again in the future.

I will contact you personally to discuss matters forward.

We hope your next stay at our hotel will be a pleasant journey and look forward to welcoming you back again.

Kind regards,

John Watson
Guest Relations Manager

Dear Sara,

Thank you for taking the time to write a review of your recent stay at our hotel. We greatly appreciate your comments.

It was concerning to read about the issues you encountered during your recent stay and I would like to apologize for the inconvenience it caused. I have personally inspected your room and it seems that we dropped the ball in a few areas during your visit.

Unfortunately during your stay, the city council had ordered a partial lockdown due to the COVID-19 pandemic. As we informed you, this implied that we were working with less than half of our staff. I have met with our whole staff to ensure that any further inconvenience will be minimized in the future.

Although we try to create the best experience possible for our hotel guests during a global pandemic, we endeavour to meet the room requirements of all our guests at all times. Thereby, we would like to <u>compensate for half</u> of your stay at our hotel and hope that you will visit us again in the future.

I will contact you personally to discuss matters forward.

We hope your next stay at our hotel will be a pleasant journey and look forward to welcoming you back again.

Kind regards,

John Watson
Guest Relations Manager

Dear Sara,

Thank you for taking the time to write a review of your recent stay at our hotel. We greatly appreciate your comments.

It was concerning to read about the issues you encountered during your recent stay and I would like to apologize for the inconvenience it caused. I have personally inspected your room and it seems that we dropped the ball in a few areas during your visit.

Unfortunately during your stay, the city council had ordered a partial lockdown due to the COVID-19 pandemic. As we informed you, this implied that we were working with less than half of our staff. I have met with our whole staff to ensure that any further inconvenience will be minimized in the future.

Although we try to create the best experience possible for our hotel guests during a global pandemic, we endeavour to meet the room requirements of all our guests at all times. Thereby, we would like to <u>apologize</u> to you for your stay at our hotel and hope that you will visit us again in the future.

I will contact you personally to discuss matters forward.

We hope your next stay at our hotel will be a pleasant journey and look forward to welcoming you back again.

Kind regards,

John Watson
Guest Relations Manager

Appendix B

Tests of Between-Subjects Effects

Dependent Variable: How likely are you to book a stay at this hotel?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	450,140[a]	34	13,239	34,329	,000	,718
Intercept	952,845	1	952,845	2470,661	,000	,843
Compensation_offered	96,186	2	48,093	124,701	,000	,352
Reviewer_expertise	21,436	4	5,359	13,895	,000	,108
Degree_of_neg	51,563	2	25,781	66,849	,000	,226
Compensation_offered * Reviewer_expertise	56,727	8	7,091	18,386	,000	,243
Compensation_offered * Degree_of_neg	,618	3	,206	,534	,659	,003
Reviewer_expertise * Degree_of_neg	30,268	6	5,045	13,081	,000	,146
Compensation_offered * Reviewer_expertise * Degree_of_neg	13,370	8	1,671	4,333	,000	,070
Error	177,020	459	,386			
Total	7759,000	494				
Corrected Total	627,160	493				

a. R Squared = ,718 (Adjusted R Squared = ,697)

Table 27: An overview of the means, three-way ANOVA output considering the interaction effects between offering compensation, reviewer expertise, degree of negativity and booking likelihood. It should be noted that (*) indicates a statistically significant result (p < 0.05).

73

Tests of Between-Subjects Effects

Dependent Variable: How would you rate this hotel? Please note that a higher number denotes a hiç

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1254,083[a]	34	36,885	27,961	,000
Intercept	2896,080	1	2896,080	2195,384	,000
Compensation_offered	245,246	2	122,623	92,955	,000
Reviewer_expertise	77,675	4	19,419	14,720	,000
Degree_of_neg	171,507	2	85,753	65,006	,000
Compensation_offered * Reviewer_expertise	159,858	8	19,982	15,148	,000
Compensation_offered * Degree_of_neg	7,485	3	2,495	1,891	,130
Reviewer_expertise * Degree_of_neg	57,614	6	9,602	7,279	,000
Compensation_offered * Reviewer_expertise * Degree_of_neg	30,497	8	3,812	2,890	,004
Error	605,498	459	1,319		
Total	23651,000	494			
Corrected Total	1859,581	493			

a. R Squared = ,674 (Adjusted R Squared = ,650)

Table 28: An overview of the means, three-way ANOVA output considering the interaction effects between offering compensation, reviewer expertise, degree of negativity and hotel evaluation score. It should be noted that (*) indicates a statistically significant result (p < 0.05).

Tests of Between-Subjects Effects

Dependent Variable: How likely are you to recommend this hotel to friends and/or family?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	451,411^a	34	13,277	36,288	,000
Intercept	927,589	1	927,589	2535,308	,000
Compensation_offered	95,653	2	47,827	130,721	,000
Reviewer_expertise	27,166	4	6,792	18,563	,000
Degree_of_neg	51,467	2	25,733	70,335	,000
Compensation_offered * Reviewer_expertise	59,050	8	7,381	20,175	,000
Compensation_offered * Degree_of_neg	1,316	3	,439	1,199	,310
Reviewer_expertise * Degree_of_neg	24,673	6	4,112	11,240	,000
Compensation_offered * Reviewer_expertise * Degree_of_neg	12,513	8	1,564	4,275	,000
Error	167,934	459	,366		
Total	7774,000	494			
Corrected Total	619,344	493			

a. R Squared = ,729 (Adjusted R Squared = ,709)

Table 29: An overview of the means, three-way ANOVA output considering the interaction effects between offering compensation, reviewer expertise, degree of negativity and recommendation to friends and/or family. It should be noted that (*) indicates a statistically significant result (p < 0.05).

Compensation	Response rate	Satisfaction	Expertise	Trustworthiness	Helpfulness
Full	3.79	3.78	4.38	4.35	4.32
Partial	2.93	3.02	4.42	4.49	4.41
None	1.67	1.66	4.22	4.24	4.25
Sig.	0.000*	0.000*	0.189	0.302	0.291
Overall	F > P > N	F > P > N	-	-	-

Table 30: An overview of the means and pairwise comparison test of the response rating and satisfaction, the reviewer's rated expertise and trustworthiness and the helpfulness of the review considering the conditions in Hypothesis 1. It should be noted that (*) indicates a statistically significant result ($p < 0.05$).

Compensation	Rate response	Satisfied	Expertise	Trustworthiness	Helpfulness
Full	4.77	4.80	4.25	4.25	4.26
Partial	4.55	4.52	4.17	4.14	4.12
None	2.85	3.12	4.32	4.41	4.39
Sig.	0.000*	0.013*	0.546	0.238	0.301
Overall	F = P > N	F = P > N	-	-	-

Table 31: An overview of the means and pairwise comparison test of the response rating and satisfaction, the reviewer's rated expertise and trustworthiness and the helpfulness of the review considering the conditions in Hypothesis 3. It should be noted that (*) indicates a statistically significant result ($p < 0.05$).

Compensation	Rate response	Satisfied	Expertise	Trustworthiness	Helpfulness
Full	4.15	4.13	1.95	1.96	1.95
Partial	3.90	3.95	2.10	2.24	2.24
None	3.00	3.41	1.98	2.00	2.05
Sig.	0.000*	0.000*	0.805	0.405	0.339
Overall	F = P > N	F = P > N	-	-	-

Table 32: An overview of the means and pairwise comparison test of the response rating and satisfaction, the reviewer's rated expertise and trustworthiness and the helpfulness of the review considering the conditions in Hypothesis 3. It should be noted that (*) indicates a statistically significant result ($p < 0.05$).

Compensation	Rate response	Satisfied	Expertise	Trustworthiness	Helpfulness
Full	4.73	4.76	2.20	2.66	2.68
Partial	4.66	4.46	2.07	2.44	2.51
None	3.01	3.72	2.02	2.43	2.40
Sig.	0.011*	0.000*	0.296	0.303	0.277
Overall	F = P > N	F = P > N	-	-	-

Table 33: An overview of the means and pairwise comparison test of the response rating and satisfaction, the reviewer's rated expertise and trustworthiness and the review's helpfulness considering the conditions in Hypothesis 4. It should be noted that (*) indicates a statistically significant result (p < 0.05).