Is webcare good for business?
A study of the effect of managerial response strategies to online reviews on hotel bookings

Ana Isabel Lopes
Department of Communication Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands and
Department of Marketing, University of Antwerp, Antwerp, Belgium
Edward C. Malthouse
Department of Integrated Marketing Communication, Northwestern University, Evanston, Illinois, USA, and
Nathalie Dens and Patrick De Pelsmacker
Department of Marketing, University of Antwerp, Antwerp, Belgium

Abstract
Purpose – Engaging in webcare, i.e. responding to online reviews, can positively affect consumer attitudes, intentions and behavior. Research is often scarce or inconsistent regarding the effects of specific webcare strategies on business performance. Therefore, this study tests whether and how several webcare strategies affect hotel bookings.

Design/methodology/approach – We apply machine learning classifiers to secondary data (webcare messages) to classify webcare variables to be included in a regression analysis looking at the effect of these strategies on hotel bookings while controlling for possible confounds such as seasonality and hotel-specific effects.

Findings – The strategies that have a positive effect on bookings are directing reviewers to a private channel, being defensive, offering compensation and having managers sign the response. Webcare strategies to be avoided are apologies, merely asking for more information, inviting customers for another visit and adding informal non-verbal cues. Strategies that do not appear to affect future bookings are expressing gratitude, personalizing and having staff members (rather than managers) sign webcare.

Practical implications – These findings help managers optimize their webcare strategy for better business results and develop automated webcare.

Originality/value – We look into several commonly used and studied webcare strategies that affect actual business outcomes, being that most previous research studies are experimental or look into a very limited set of strategies.

Keywords Webcare, Managerial responses, eWOM, Online reviews, Hotel bookings

Paper type Research paper

© Ana Isabel Lopes, Edward C. Malthouse, Nathalie Dens and Patrick De Pelsmacker. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/4.0/legalcode

The authors would like to thank Mr. Christian Holthof for his help in getting access to the data.
Funding: The authors gratefully acknowledge financial support from the University of Antwerp BOF research fund (UAntwerpen DOCPRO FP170222).
1. Introduction
Electronic word-of-mouth (eWOM) has rapidly gained importance in consumer decision-making. Online reviews, a form of eWOM in which users or experts evaluate products or services based on their experience, are critical in the service industry (e.g. Dens et al., 2015; Ruiz-Mafe et al., 2020). As the volume of eWOM increases constantly, organizations should manage online reviews to benefit their business (Tsiotsou and Diehl, 2022). Webcare – organizational responses to social media messages – is an essential tool to mitigate the negative influence of negative eWOM and boost positive eWOM. But as its effectiveness depends on many factors (see Dens et al., 2015; Jacobs and Liebrecht, 2023), organizations struggle with whether and how to respond. The present study addresses the following research question: Does providing webcare (vs not responding) influence future hotel bookings, and how do specific webcare strategies affect bookings? A webcare strategy refers to the type of reply (e.g. timeliness, style or content, see section 2.1 and Figure 1 for more detail). The strategies included in this study are based on the classification by Lopes et al. (2023), namely signing with name and function, timeliness, tailoring, tone of voice, redirect customers to a private channel, request more information, use of apologies, compensation and defensiveness. While the effect of some strategies is well-documented (e.g. apologizing), others are barely explored in previous research. For example, in the current study, we explore the effects of inviting customers for another visit or showing gratitude on business performance, which have not yet been studied, despite these being commonly used webcare strategies.

A recent literature review (Lopes et al., 2023) shows that there is little consensus as to what response strategies benefit either brand-related (e.g. brand trust or purchase intention) or commercial (e.g. sales or competitive performance) outcomes (or both). For instance, Xie et al. (2017b) find a negative effect of executives replying on financial performance, while Xie et al. (2017a) find the opposite. Purani and Jeesha (2023) document that responding to all reviews is better for review readers’ customer engagement intentions, while Anderson and Han (2016) find that responding to more than 85% of reviews results in lower revenues than not responding at all, and Xie et al. (2016) find no impact of the response rate on hotel performance.

![Figure 1. Expected effects of the webcare strategies on hotel bookings](image-url)
This lack of consensus furthers the need for more research on how response strategies affect business performance. Most previous research consists of (lab) experiments with limited external validity, studying outcomes as trust or purchase intentions that do not necessarily predict actual behavior (Morwitz et al., 2007). To fill these gaps, the current study draws from justice theory (Tax and Brown, 1998; Ghosh and Mandal, 2020) to investigate how frequently used webcare strategies affect future bookings as a relevant commercial outcome. Previous research shows that successful webcare depends on consumers’ justice perceptions (e.g. Ghosh and Mandal, 2020; Javornik et al., 2020; Dens et al., 2015). These perceptions reflect the extent to which customers perceive their treatment as fair, considering how they are treated compared to others as well as the severity of the failure. Justice perceptions are typically classified as distributive, interactional or procedural (Ghosh and Mandal, 2020). Distributive justice reflects the perceived fairness of the offered resolution in offsetting the loss experienced by the service failure (Tax and Brown, 1998). In the context of webcare, distributive justice could be achieved through monetary (i.e. a discount) or psychological (i.e. an apology) compensation (Gelbrich and Roschk, 2011). Interactional justice relates to perceptions of how customers are treated during service recovery (Tax and Brown, 1998). Providing explanations or an empathetic tone helps realize interactional justice (Gelbrich and Roschk, 2011). Procedural justice refers to the process of complaint handling, for instance, how timely a complaint is addressed (Gelbrich and Roschk, 2011), whether or not one directs consumers to another channel or who is in charge of replying. Given the lack of research on the effect of webcare strategies (responses to online reviews) on future hotel bookings, we use justice theory as the overarching framework allowing us to derive hypotheses on the effects of webcare.

This article investigates the actual bookings received through Booking.com for a convenience sample of seven Belgian hotels over four years and reports the effect of several webcare strategies on these bookings. The study contributes to existing knowledge on webcare in several ways. First, it is one of few studies to analyze the effect of specific webcare strategies on actual bookings rather than mere customer perceptions. For many strategies under study, the effect on business performance is either under-researched or prior research documents contradictory findings. Second, as shown in Table 1, this is the first study to our knowledge that investigates a comprehensive set of webcare strategies in one parsimonious model, while controlling for both bookings’ seasonality, the specific hotel, and previous bookings, which benefits the generalizability of our findings. Third, we develop an automated machine-learning approach for coding webcare responses and test the performance of different machine-learning text classifiers.

In the following sections we survey previous literature on the effects of the webcare strategies identified by Lopes et al. (2023) and derive hypotheses for the effects of webcare strategies on future hotel bookings based on justice theory. Studies focusing on business performance (such as bookings) are scarce. Therefore, we must derive our hypotheses based on studies on the effects of other variables that can be considered proxies to hotel bookings. First, we approach the webcare strategies that might be adequate to manage online reviews regardless of their valence; next, we investigate webcare strategies specifically for negative reviews.

2. Theoretical development

2.1 Effects of webcare strategies

In what follows, we distinguish between managerial responses to online reviews, regardless of whether they are positive or negative. However, responding to negative reviews may require specific webcare strategies to elicit justice perceptions. Webcare strategies specific to negative reviews are typically classified as accommodative – complaisant and comprising corrective action, compensation, and/or mortification – or defensive – denial and evasion of
responsibility (Lopes et al., 2023; Einwiller and Steilen, 2015). Based on justice theory, both types can, in principle, elicit justice perceptions, since they acknowledge customers in their concerns, even if the claims may be refuted in the case of defensive webcare.

Most previous research finds a positive effect of providing webcare on subsequent rating (e.g. Wang and Chaudhry, 2018; Proserpio and Zervas, 2017; Ravichandran and Deng, 2023), subsequent review volume (e.g. Proserpio and Zervas, 2017), trust (Bhandari and Rodgers, 2018), consumer sentiment (Ma et al., 2015), satisfaction (Zhao et al., 2020), attitudes and behavior (Le and Ha, 2021), and hotel revenue (Anderson and Han, 2016). In contrast, Xie et al. (2014) report that responding (vs not responding) to online reviews harms hotels’ RevPAR (Revenue per available room). Bhandari and Rodgers (2018) also find a negative effect of responding to negative reviews on purchase intentions, although this effect is partly offset by an indirect positive effect through brand trust. They argue that webcare helps to reinforce the brand’s promise to deliver a product of value. Replying to reviews signals interactional and procedural justice and can be a mean to offer distributive justice (Gelbrich and Roschk, 2011; Smith et al., 1999). Providing webcare can, therefore, restore justice for reviewers and reassure bystanders who see the review and response.

H1. Providing webcare (versus not) positively influences hotel bookings.

In practice, managers who decide to engage in webcare face many choices. The first is who should respond? Xie et al. (2017b) find that webcare provided by hotel executives lowers future business performance compared to webcare provided by the staff because staff’s

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Li et al. (2018)</th>
<th>Xie et al. (2017b)</th>
<th>Xie et al. (2017a)</th>
<th>Current study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotel Sales Revenue</td>
<td>Financial performance (revenue, average daily rate, and occupancy)</td>
<td>RevPAR</td>
<td>Future hotel bookings</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control variable(s)</th>
<th>Seasonality</th>
<th>Hotel characteristics</th>
<th>Hotel characteristics</th>
<th>Seasonality, hotel, previous bookings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webcare strategies</td>
<td>Manager’s name and function</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Timeliness</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tailoring (*)</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tone of voice</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Redirect to private channel</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Request more information</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Express gratitude</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Apology (*)</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Compensation (*)</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Defensiveness</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Note(s): *Li et al. (2018) label responses including apology, problem acknowledgment, acceptance of responsibility or remedy attempts as an accommodative response. As they do not shed light on the individual effects, their results for those strategies are not comparable to the present study. The authors also claim to study tailoring; however, this is operationalized as offering either accommodative or defensive webcare, not fitting the designation of tailoring adopted in the current work.

Source(s): Developed by the authors

Table 1. Previous research using business outcomes as dependent variable versus the current study
operational insights enable them to address consumer comments in a more relevant and helpful way. Kniesel et al. (2016) find no significant difference in bystanders’ attitudes between managers’ and staff members’ answers.

Tathagata and Amar (2018) find that webcare responding to negative reviews leads to higher forgiveness from bystanders when it contains high ownership (webcare provided by an individual with personal details such as name and title, e.g. owner, manager) as compared with being provided by a team or department. Similarly, Xie et al. (2017a) find that a higher portion of executive responses (i.e. by managers) benefits hotel performance. By seeing a name, consumers can attribute their blame to this identifiable person rather than the firm (Tathagata and Amar, 2018). Webcare displaying ownership also contributes to the credibility of the reply (Davidow, 2000), leading to greater perceptions of interactional justice. Webcare ownership can further signal efficient organizational procedures (i.e. structures a company has in place to provide a smooth complaint-handling process) (Gelbrich and Roschk, 2011), impacting perceived procedural justice. In sum, webcare signed with the manager’s name and function should benefit both interactional and procedural justice perceptions of bystanders.

H2. The presence of a manager’s name and function (versus staff members or hotel name) in webcare positively influences hotel bookings.

An important and widely studied webcare strategy is timeliness: how soon to respond? A content analysis conducted by Mate et al. (2019) finds that the majority of customer reviews are responded to within four days of being received, which is what the authors consider timely. Previous findings seem consistent: giving a timely (versus late) response increases readers’ trust, decreases their concern (Sparks et al., 2016), and leads to higher levels of readers’ forgiveness (Ghosh, 2017) and more satisfaction (Zhao et al., 2020). Timely responses also increase future review volume (Sheng, 2019), improve the valence of future reviews (Sheng et al., 2021), and benefit financial performance (Xie et al., 2017a, b). Based on justice theory, it is expected that a timely reply will lead to higher perceptions of procedural justice, which in turn increases satisfaction with complaint handling (Gelbrich and Roschk, 2011).

H3. Providing timely (versus late) webcare positively influences hotel bookings.

Regarding the content of the response, managers are faced with the question of which stylistic elements to use? Tailoring webcare, where webcare is adapted to the content of the review, is commonly studied. Moreover, Min et al. (2015) and Palese et al. (2021) find that paraphrasing a complaint (a form of tailoring) in response to a negative review causes potential guests to evaluate the response more favorably than a generic answer. Darani et al. (2023) find that mimicry (tailoring by using similar words to the ones used in the review) increases the star ratings of subsequent reviews and Jin et al. (2023) find that, when directed at negative reviews, personalizing a message increases perceived response helpfulness.

Providing explanations is a form of tailoring since they typically refer to what is mentioned in the review. Similar to tailoring, previous research finds that webcare directed at negative reviews containing strong explanations can produce high consumer forgiveness (Ghosh, 2017). Xie et al. (2017a) find no significant relationship between the match rate between text in the review and the reply (i.e. tailoring) and hotel performance. Another study on how tailoring affects business performance finds a negative effect of simply repeating the topics mentioned in the review on hotel financial performance (Xie et al., 2017b). This finding reinforces the idea that explanations are an important component of webcare because simply repeating what is mentioned in the review might lead to negative outcomes. Gelbrich and Roschk (2011) state that favorable employee behavior, which comprises helping the complainant to understand why a failure occurred, is a strong predictor of interactional justice. By signaling an effort from the organization to explain how the issue expressed in the review is addressed, providing explanations could trigger perceptions of interactional justice.
Providing tailored wecare (versus generic) positively influences hotel bookings. Another stylistic element is the tone of voice. A conversational human tone is achieved by personalizing responses (i.e. including the reviewer’s name), inviting guests to visit again or using non-verbal cues (e.g. abbreviations, emoticons, words in upper case) (Liebrecht et al., 2021). Compared to a professional tone, using a conversational tone makes bystanders less concerned about the problem expressed in the review (Sparks et al., 2016). Barcelos et al. (2018) find that a conversational versus corporate tone of voice increases a consumer’s purchase intentions. Similar to Gelbrich and Roschk (2011)’s framework, Javornik et al. (2020) show that using a conversational tone leads to more positive observer perceptions of complaint handling than a corporate voice by positively influencing interactional justice perceptions. Jacobs and Liebrecht (2023) find that using a conversational human voice (versus a corporate voice) positively influences all three justice dimensions.

Using elements suggesting a conversational human tone of voice (versus a corporate tone), namely personalizing, inviting guests for another visit, and using non-verbal cues, positively influences hotel bookings.

2.2 Effects of wecare strategies responding to negative reviews

Finally, while the aforementioned strategies apply to both positive and negative reviews, much of the wecare literature pertains to how to respond to negative reviews. As we review below, there have been many studies of specific wecare strategies, especially accommodative ones, including changing to a private channel, inquiring for further information, expressing gratitude, apologizing, and offering compensation.

In which channel to respond? Previous research suggests that firms should publicly contact dissatisfied reviewers and invite them to engage in a private conversation, thereby changing the channel in which the conversation occurs (Zhang et al., 2019). Initial public replies show bystanders that the organization cares about dissatisfied reviewers and is prepared to solve the issues mentioned in the review, which signals interactional justice. Steering the conversation away from the public eye could help prevent the negative information from spreading virally or creating online firestorms (Zhang et al., 2019), especially in the early stages of discussion (Herhausen et al., 2019). Online firestorms can threaten an organization’s reputation and financial outcomes. By showing a willingness to work on solving the complaint, inviting reviewers to follow up in a private channel should elicit perceptions of procedural and interactional justice from bystanders.

Inviting reviewers to follow up on a review in a private channel (versus staying in a public channel) positively influences hotel bookings.

Previous research finds that inquiring for further information is the most common strategy in practice, but does not improve satisfaction with how the complaint was handled (Einwiller and Steilen, 2015). Previous research on service failure recovery points out that initiating the process for service recovery, such as by asking for more information, can lead to higher perceptions of interactional justice, which leads to satisfaction with the service encounter (Smith et al., 1999) and should translate to business outcomes.

Asking for more information (versus not) positively influences hotel bookings.

Expressing gratitude is the second most common wecare strategy (Sparks and Bradley, 2017). Previous research shows that using words such as “thank you” improves satisfaction with complaint handling (Einwiller and Steilen, 2015). In the same vein, Farias et al. (2022) find that publicly expressing gratitude supports an increase in business reputation. Resorting to justice theory, we can argue that showing gratitude is a form of displaying attentiveness,
which according to (Gelbrich and Roschk, 2011) is a form of favorable employee behavior. In turn, this leads to higher perceptions of interactional justice. Based on this, we expect that:

H8. Showing gratitude (versus not) positively influences hotel bookings.

A frequently used and studied accommodative strategy is apologizing (e.g. Zhang and Vásquez, 2014; van Hooijdonk and Liebrecht, 2021). Previous research finds that bystanders who see wecare containing an apology (versus none) have lower behavioral intentions (Kim et al., 2016). However, most research on apologizing (versus not) does not find significant effects or even a negative influence on attitudes and intentions. Dens et al. (2015) find that apologizing does not significantly raise readers’ attitudes or patronage intentions compared to no response, even when most reviews are positive. Similarly, van Hooijdonk and Liebrecht (2021) find that the presence of an apology (versus absence) does not enhance brand reputation. By apologizing, firms assume guilt in a reviewer’s accusations (Lee and Song, 2010). However, at the same time, apologizing is seen as representing a caring attitude and showing compassion for the experienced failure (Mate et al., 2019). Previous research on service recovery shows that apologizing is strongly linked to customers’ perceptions of interactional justice (Smith et al., 1999; Wirtz and Mattila, 2004).

H9. Using apologies (versus not) has a positive effect on hotel bookings.

Sparks and Bradley (2017) note that one of the most effective responses after a negative review is to offer compensation (e.g. a discount on a future purchase). Liu et al. (2019) suggest compensation as the optimal response for less severe failures. Offering compensation also mitigates the virality (how much a message spreads online) of negative eWOM when used in evolved stages of online firestorms (Herhausen et al., 2019). Compensation benefits readers’ purchase intention and brand perceptions (Treviño and Castaño, 2013) as well as brand reputation (Rose and Blodgett, 2016). Smith et al. (1999) state that offering compensation after a service failure leads to higher perceptions of distributive justice, increased satisfaction with the service encounter, and more positive behavioral intentions (Gelbrich and Roschk, 2011).

H10. Providing wecare that does offer compensation (versus not) positively influences hotel bookings.

Previous research finds that defensive responses strengthen bystanders’ perceptions that the company was at fault rather than “no action” (Lee and Song, 2010). Other studies, however, do not find a significant negative effect of defensive responses compared to no response. Weitzl and Hutzinger (2017) find that the only defensive strategy that significantly adversely impacts failure attribution is vouching (i.e. countering negative comments with favorable statements), trivializing the review claims or expressing doubts do not. Credible, defensive responses might also strengthen bystander-brand relationships (Weitzl and Hutzinger, 2019).

Moreover, “flyting”, a ritualized exchange of insults between two or more interlocutors, could help brands bolster their ideological positioning by engaging with opposing sides in verbal contests (Scholz and Smith, 2019). Lee and Cranage (2014) find that when there is little consensus among negative reviews (meaning that some reviews are positive), defensive responses are more effective than no response to prevent a negative bystander attitude. Treviño and Castaño (2013) find that hotels providing wecare, even if defensive, are perceived as giving more importance to customer service than hotels that do not respond. This attention paid to customers signals that the organization has procedures in place to satisfactorily address complaints. Based on Gelbrich and Roschk (2011), it can be argued that providing defensive wecare facilitates perceptions of procedural justice. Furthermore, Li et al. (2018) find that, when a hotel adopts a defensive response to an ordinary negative review (reflecting dislike, mismatched preferences, unrealistic expectations or occasionally...
unreasonableness on the part of the reviewer), its revenue increases; there is only a negative
effect of defensiveness in case of product failure reviews.

H11. Using defensive webcare (versus not) positively influences hotel bookings.

Figure 1 shows the expected effects of the studied webcare strategies on hotel bookings.

3. Methodology
We use text mining and panel data models to study how different webcare strategies affect
hotel bookings by combining online reviews and their managerial responses on Booking.com
with booking data through Booking.com from seven Belgian hotels for almost four years
(February 2016 until December 2019). We asked 66 hotels for booking data, and seven agreed
to participate. The seven hotels include five chain hotels and two independent ones. They are
diverse in size (10–150 rooms), with different target audiences (three holiday, two business,
and two business/holiday). On Booking.com, the managerial responses are displayed under
each review and thus are noticeable to bystanders. The online reviews and managerial
responses were supplied by a company that provides hotels with integrated tools and
processes to manage guest satisfaction. The hotels provided the number of bookings they
received each week from Booking.com. In sum, for each hotel, we know what reviews were
posted (and when), the managerial responses (if any) provided, and the number of weekly
bookings received through Booking.com. In total 18,320 reviews were posted on Booking.com
during the mentioned period, with 5,564 receiving managerial responses.

3.1 Description and operationalization of webcare variables
Creating variables to measure the use of different webcare strategies required several steps.
First, all reviews were translated into English with the Google Translate API. Second, we
determined how to measure all webcare variables discussed in the literature review. Table A1
of appendix A summarizes the variables used in this study, including a short definition and
the chosen method for each one (more details on the method are provided next). More details
for each variable can be found in Appendix A. Some variables could be automatically coded
or coded using an existing dictionary approach. When this was not the case, we further
developed a classifier, as explained in the next section.

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

3.1.1 Manual coding. Due to the large volume of responses and the study’s goal of
developing automated approaches, a training sample was coded with human coders. We then
applied leading machine classifiers to determine which one worked best for classifying
webcare responses so that all responses could be coded. Details are provided here, Appendix
A section 1. In total, 810 responses were coded, which is about 15% of the 5,564 cases.

3.1.2 Performance evaluation of machine learning classifiers. After coding the training
sample, the next step was to evaluate the performance of leading machine learning classifiers
including support vector machines (SVM), boosted trees (GBM), random forests (RF), and
naive Bayes (NB) (for details about these methods see James et al., 2021). We also tested the
bidirectional encoder representations from transformers (BERT) pre-trained model (Devlin
et al., 2018), which is consistently one of the top algorithms for test classification (Cunha et al.,
2021). Details of the methodology are provided in Appendix B. The first step was to determine
which classifier performs best for each variable. A single test set was set aside consisting of
10% of the coded data (82 cases) and each model was trained on the remaining cases. Having
estimated the models and tuned regularization parameters we applied the models to the test
set to evaluate accuracy. As shown in Table B1 (appendix B), for all variables, BERT either
has the best accuracy or has accuracy close to the best. BERT will be used to estimate the variables used in the study’s panel models.

3.1.3 Application of BERT to text classification. Having established that BERT is the best approach, 10-fold cross-validation was used with all 810 training cases to compute out-of-sample accuracy (percentage of correctly classified cases) and AUC (area under a ROC curve) measures. Performance metrics for BERT are provided in Table 2. Accuracy and AUC are highly correlated and tell a consistent story. Most dimensions could be classified with near-perfect accuracy, but BERT had more difficulty classifying the more subjective dimensions of explanations, invitations, and defensiveness. In the worst cases (defensiveness and invitations), BERT still correctly classifies at least 85% of the cases with AUC values greater than 80%.

Having trained the BERT text classifier, it was applied to the whole sample of 5,564 responses. The next step was to aggregate the data by week and match it with the booking data from the next week. Using weekly data avoids within-week seasonality issues. The data were aggregated by summing all the instances in each week where a certain strategy was used by the hotel. In total, there are 192 weeks of data for the seven hotels, totalizing 1,344 observations.

3.2 Description of the model

This section describes the model used to study how the webcare variables are associated with bookings during the next week.

3.2.1 Assessing the correlation among the independent variables. The first step was to study relationships among predictor variables to understand where multicollinearity could be problematic. Five webcare variables measure who signed the response. Some had very low occurrences (e.g. signed by the department). Principal component analysis (see Appendix C) showed that these variables load on three different factors. Subsequently, they were combined into three new variables: signed by staff or department (sigDepStaf), signed by manager or with name (sigNameMgr), and signed by hotel (sigHotel).

A correlation matrix (see Appendix C, Table C1) showed high correlations among some variables ($r > 0.81$), namely between tailoring and defensiveness, invite for a visit, explanation, and personalization, and between defensiveness and explanation, explanation and invite for a visit, and gratitude and invite for a visit. Large correlations cause high variance inflation factors. We explain how correlation issues were addressed in the next section.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signed with department</td>
<td>0.96</td>
<td>0.91</td>
</tr>
<tr>
<td>Signed with hotel</td>
<td>0.99</td>
<td>0.95</td>
</tr>
<tr>
<td>Signed by manager</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>Signed with name</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Signed by staff</td>
<td>0.98</td>
<td>0.93</td>
</tr>
<tr>
<td>Tailoring</td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td>Explanation</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>Invitation for visit</td>
<td>0.86</td>
<td>0.82</td>
</tr>
<tr>
<td>Non-verbal cues</td>
<td>0.98</td>
<td>–</td>
</tr>
<tr>
<td>Defensiveness</td>
<td>0.93</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Note(s): *For the variable defensiveness, a balanced training set was used (same number of cases for each label)

Source(s): Developed by the authors

Table 2. Cross-validated accuracy and AUC scores achieved by BERT
3.2.2 Creating a parsimonious regression model. A panel Poisson regression model was created in which the number of weekly bookings was the dependent variable (nextbook). Poisson regression models are often used when the dependent variable is a count (Neter et al., 1996, section 14.13). The model controls for seasonality (time), bookings the week before (book), and idiosyncratic hotel characteristics by including six 0–1 dummy variables (hotel). This panel design is robust to many threats to internal validity. The time variable should account for outside events that could affect bookings across hotels, such as a festival or holiday. Hotel dummies account for systematic differences between the hotels, e.g. unobserved causal factors such as one hotel having more mass advertising than others. Including lagged bookings (book) further controls for these threats [1].

The model also includes a dummy variable for the webcare treatment that takes the value 0 when there are no hotel responses and 1 when the hotel responds to at least one review during a given week (webcare). A log transformation was performed on the variables nextbook, book, defensive, invitevisit, nonverbal, apology, compensate, change, gratitude, information, personal, sigDepStaf, and sigNameMgr since these are right skewed with very few observations in the tail.

The first estimated model includes all variables (Appendix C, Table C4) and has severe multicollinearity issues, with VIF >10 for the variables tailor, explain, personal, invitevisit, and sigNameMgr. Since tailor and explain were highly correlated with the other variables (VIFs of 24 and 11, respectively), they were dropped from the analysis. The sigNameMgr has VIF = 17.92 because it correlated with sighotel; dropping the latter reduced sigNameMgr’s VIF to 5.89.

Table C5 (Appendix C) presents a summary of the variables included in the final model. The default log link function for a Poisson model was used, and therefore the model predicts the log-mean number of bookings. The above model was fitted to test how several webcare strategies influenced future bookings. Compared to the previous model that included all webcare variables, this model had all VIF values less than 10, although some are greater than five.

4. Results

Basic descriptive statistics show that inviting for another visit is the most often used strategy (3,686/5,564 responses), followed by personalizing the response with the name of the reviewer (n = 2,678) and by showing gratitude (n = 2,093). There is a similar number of instances where hotels were defensive (n = 1,477) and apologetic (n = 1,102). Less often used strategies include offering compensation (n = 510), asking for more information (n = 418) and using non-verbal cues (n = 169). There are 2,045 signatures from the department or the staff, and 5,715 signatures that include a name and/or come from the manager.

Table 3 presents the results for the final model, which has a McFadden’s Pseudo $R^2 = 1-(8276/63486) = 87.0\%$. To evaluate the robustness of the model, a lasso regression (James et al., 2021, section 6.2) is also fitted with the same variables selecting the shrinkage hyperparameter with cross-validation. Comparing the estimates of the GLM and lasso models shows that the effect sizes are robust across models. This indicates the robustness of the model since in linear models there is no penalization for the model’s choice of weights, which could lead to overfitting.

We focus our discussion on interpreting coefficients related to our hypotheses. Responding to reviews at least once in a given week has a positive effect ($b = 0.055$, $z = 5.2$, $p < 0.0001$) on the number of bookings received the next week, confirming H1. Because the model predicts the log expected bookings, responding to reviews is associated with an increase of $e^{0.055} = 1.056$ (5.6%) in bookings. Answering H2, managers signing with their name or function (SigNameMgr) positively influences future bookings ($b = 0.02$, $z = 4.0$, $p < 0.0001$).
**Table 3.** Results for regression model (GLM)

| Coefficients GLM | Estimate | Std. error | z value | Pr(>|z|) | VIF GLM | Lasso Estimate |
|------------------|----------|------------|---------|----------|---------|----------------|
| (Intercept)      | 2.13     | 0.06       | 37.17   | <2e-16 *** |        | 2.02           |
| hotel_id = 2     | -0.08    | 0.01       | -7.13   | 9.86e-13 *** | 43.03   | -0.06          |
| hotel_id = 3     | -1.18    | 0.03       | -37.70  | <2e-16 *** |        | -1.09          |
| hotel_id = 4     | -0.84    | 0.02       | -36.38  | <2e-16 *** |        | -0.77          |
| hotel_id = 5     | -0.67    | 0.02       | -35.41  | <2e-16 *** |        | -0.61          |
| hotel_id = 6     | -0.01    | 0.01       | -0.63   | 0.53      |         | 0.0002         |
| hotel_id = 7     | -0.58    | 0.02       | -27.50  | <2e-16 *** |        | -0.54          |
| Book             | 0.61     | 0.008      | 78.70   | <2e-16 *** | 4.82    | 0.63           |

Note(s): Signif. codes: 0 “***” 0.001 “**” 0.01 “*” 0.05
Null deviance: 120,407 on 1,143 degrees of freedom
Residual deviance: 9,988 on 1,119 degrees of freedom
Source(s): Developed by the authors

*p < 0.0001, e^{0.02} = 1.202*, while having staff members or departments signing (sigDepStaff) is not significantly associated with future bookings (*p > 0.05*).

H3 (timeliness) and H4 (tailoring) could not be tested because they had to be dropped from the analysis. To test H5, we analyzed three variables pertaining to the tone of voice. Nonverbal cues (*b = -0.05, z = -3.7, p < 0.0001*) and an invitation for another visit (invitevisit) (*b = -0.02, z = -3.3, p = 0.001*) have significant negative associations with bookings, representing a decrease of 4.9% (*e^{-0.05} = 0.9512*) and 2% (*e^{-0.02} = 0.9802*) respectively, in the number of bookings the following week. The use of personalization (personal) did not significantly affect future bookings (*p > 0.05*). Therefore, H5 stating that a conversational tone positively influences hotel bookings is not confirmed.

Asking to change to a private channel positively influences future bookings (change, *b = 0.04, z = 4.0, p < 0.0001*), as expected in H6, causing an increase of 4.1% (*e^{0.04} = 1.040*) in the bookings received in the following week. Regarding H7, the effect of asking for more information (information) (*b = -0.04, z = -4.3, p < 0.0001*) is negatively associated with the number of bookings that hotels receive the following week. This means that asking for more information lowers hotel bookings by 3.9% (*e^{-0.04} = 0.9607*). Expressing gratitude does not significantly affect future hotel bookings (*p > 0.05*), which does not confirm H8. Unlike what was predicted in H9, the results show that the use of apology (*b = -0.05, z = -6.3, p < 0.0001*) has a negative effect on future bookings. Apologizing lowers future hotel bookings by 4.9% (*e^{-0.05} = 0.9512*). As predicted in H10, offering compensation significantly increases future bookings (*b = 0.02, z = 2.3, p = 0.04*) by 2% (*e^{0.02} = 1.020*). Finally, as predicted in H11, being defensive positively influences future bookings (*b = 0.02, z = 3.1, p = 0.01*), increasing the number of bookings the following week by 2% (*e^{0.02} = 1.020*). Figure 2 summarizes the expected effects and observed results.
5. Discussion

This study investigates the effects of common webcare strategies on future hotel bookings. The analyses showed, in descending order of Z-statistics, that responding, having managers sign the response, directing reviewers to a private channel, being defensive, and offering compensation positively affect bookings. Webcare strategies to avoid include apologizing, asking for more information, adding informal non-verbal cues such as emojis, and inviting customers for another visit. Strategies that did not have significant associations with future bookings were expressing gratitude, personalizing, and having staff members (rather than managers) sign webcare.

5.1 Theoretical contributions

This study clearly demonstrates that above and beyond the effect of merely replying, who replies, what is said and done, and how it is said matters to bystanders. This could explain why previous research has documented opposing or no effects of replying since they did not consider the content, tone or sender of the replies. Webcare enhances perceptions of service quality with potential customers, leading to an increase in sales. This study further supports justice theory (Gelbrich and Roschk, 2011; Smith et al., 1999). While our study does not measure justice perceptions as such, many of the webcare strategies that have previously been shown to determine justice perceptions, are now demonstrated to also affect bookings. Therefore, this lends support to the assumption that interactional, procedural, and distributive justice perceptions are key psychological mechanisms underlaying the effects of webcare on business performance.

We now discuss the effects further, starting with the positive effects. The rationales for effects were thoroughly developed and connected to previous research in the literature review section, and, to avoid duplication, this discussion will focus on where there are contradictions in extant research. While the finding that responding increases future bookings is consistent with many previous studies (most used a proxy for bookings) (e.g. Chen et al., 2019), this result contradicts some previous research finding a negative effect (Bhandari and Rodgers, 2018; Xie et al., 2014). These prior studies have typically not controlled for possible confounds or

---

**Figure 2.** Expected and found effects for the webcare variables on future hotel bookings

| Source(s): Developed by the authors |

<table>
<thead>
<tr>
<th>Predicted effect</th>
<th>Effect found</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>NS</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Figure 2.*

**Hotel Bookings**
considered the content of the webcare provided. Thus, we addressed the call from previous research to study how the effect of providing webcare might depend on the content of the response (van Noort et al., 2015).

The finding that using defensive webcare increases bookings defies the fairly established idea that accommodative webcare, exclusively, is the only type that can result in positive outcomes (e.g. Casado-Díaz et al., 2020). Together with other studies suggesting that organizations should in certain cases adopt defensive strategies, for instance in reply to ordinary negative reviews (Li et al., 2018), the findings point to a new perspective on webcare. This is consistent with previous findings that, by being accommodative, firms accept the blame for negative aspects described in the review, with adverse downstream effects (Lee and Song, 2010). While review writers naturally tend to make internal attributions and blame the service firms involved, review readers act as third parties. They also take other cues into account when making their judgments, such as their impression based on hotel pictures, reviewer characteristics, other situational factors, and webcare. Previous research (e.g. Dens et al., 2015) shows that the degree of consensus in a review set might also determine if accommodative or defensive strategies are preferable. If a hotel receives many negative reviews, readers are more likely to infer that the hotel does indeed possess flaws, and defensive responses will not seem credible. Reviews are often mostly positive, and bystanders are thus confronted with predominantly consistent positive reviews. Being defensive to one or a few negative reviews can, in such a case, lead to positive results. Moreover, there are occasions when customers knowingly and incorrectly report service failures or make illegitimate complaints. In these cases, companies should be able to refute such dishonest complaints. Smith et al. (1999) also suggest that the effect of recovery strategies on satisfaction depends on the magnitude and type of the failure.

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

Inviting customers for another visit and using non-verbal cues, two elements that determine the tone of voice used in the webcare message, also exert adverse effects on future hotel bookings. These findings, together with the non-significant effect found for personalizing the response with the name of the reviewer, point to a possible negative influence of using a conversational tone over a professional tone. Inviting the reviewer for another visit might be interpreted as a strategy to sell, undermining the notion that hotels respond because they care about their customers. Moreover, inviting guests to visit again is such a frequently used webcare move (Zhang and Vásquez, 2014) that bystanders might see it as meaningless, leading to a negative association with its use. The negative effect of non-verbal cues might be explained by previous research that finds that non-verbal cues used in marketer-consumer interactions, such as smiles, convey friendliness and increase perceived warmth but decrease perceived competence (Wang et al., 2017).
Finally, personalizing the response with the name of the reviewer, expressing gratitude, and having staff members sign the webcare do not have a significant effect on future bookings. Similarly to inviting for another visit, thanking the reviewer for their comment does not foster positive outcomes considering how commonly it is used (Zhang and Vásquez, 2014). Having staff members sign the messages does not lead to more bookings, as it happens when managers sign, showing that ownership of senior management is important in how bystanders have their booking decisions influenced by webcare (Tathagata and Amar, 2018). Our findings solve the contradictory effects found by Xie et al. (2017a, b) by showing that having managers reply to the reviews increases bookings.

Our findings contribute beyond extant research on webcare strategies by using real data and commercial outcomes as the dependent variable in this study. Only a small number of studies have done so (see Table 1), and they focused on a limited number of strategies, which leaves many commonly used strategies outside of the academic scope of webcare. Therefore, we are the first to show the effects of tone of voice, inviting for another visit, asking for more information, showing gratitude, and apologizing on commercial outcomes. Furthermore, we solve inconsistencies in previous studies, by showing that having managers replying to reviews increases bookings, which supports the findings by Xie et al. (2017a). The different findings from Xie et al. (2017b) could possibly be explained by the operationalization of this variable. While we have resorted to reliable and valid manual annotation to create this variable, we believe that other approaches could lead to responses written by managers being identified otherwise. For instance, a reply by a manager who occasionally signs only with a name can be classified as a reply by a staff member, especially if this classification is automatic. Given the lack of detail on the operationalization of the variables by Xie et al. (2017b), we can assume that this is a valid reason to explain the discrepancies.

5.2 Managerial implications
This study helps managers within the service industry optimize their webcare strategy for better business results. First and foremost, the act of replying positively impacts bookings and is therefore highly recommended. Organizations should employ social listening to identify online conversations on their brand and actively participate in those conversations. Second, managers should address reviews themselves or allow customer service employees to sign on their behalf. Third, managers should ensure that their response conveys responsiveness and empathy while avoiding (mere) apologies, blatant invitations for a future visit, and non-verbal cues such as exclamation marks or emojis. In the event of a negative online review, there are two possible approaches that will have a positive impact on business results. When the complaint is (or seems) legitimate, hotels should offer compensation and/or direct the complainer to a private channel. Considering that compensations can be costly for businesses (Liu et al., 2019), managers who cannot offer them consistently should consider first directing the reviewer to another channel and then offering compensation privately. Such a strategy would avoid future guests’ disappointment if they do not receive a similar offer, hindering their perceptions of distributive justice (Smith et al., 1999). However, when there is ground for the premises mentioned in the reviews to be refuted, managers should defend their perspective and present alternative facts to what is mentioned in the review.

Note that the above-described results are additive. This means that it would be most beneficial for managers to combine different strategies. A single reply could refute certain unwarranted claims and, at the same time, publicly offer compensation for a smaller issue, directing the reviewer towards a private channel for follow-up. Managerial replies are not only directed at the original reviewer and offer the potential for positive engagement or service recovery with existing customers. Perhaps more importantly, they provide strong signals of service quality to bystanders contemplating whether they should book this hotel.
among many alternatives. The words of Lord Hewart in 1924, Chief Justice of England in the case of Rex v. Sussex Justices, apply to webcare: “Justice must not only be done, but should manifestly and undoubtedly be seen to be done”.

5.3 Methodological contributions
This paper contributes methodologically by testing several machine learning classifiers to identify webcare strategies. By documenting the steps taken to create each variable, the foundations are laid for future research that faces similar challenges. Often, webcare researchers have to resort to machine learning or big data approaches to fulfill the goals of their studies (Line et al., 2020). However, the panoply of classifiers and technical options to deal with a large volume of data can seem overwhelming without proper guidance and benchmarks. Our study shows that BERT performs better than the other classifiers in labeling this type of online data, consistent with previous studies using text classifiers for similar tasks (Cunha et al., 2021).

These insights can also help in developing automated webcare, for instance, large language model–based chatbots (Koc et al., 2023). With the large volume of reviews, businesses are exploring automated ways to reply. Our methodological approach and evaluation of different machine learning classifiers to discern webcare responses is valuable to those developing automated responses. Integrating the insights from our results together with the capacities of large language models can help businesses to appropriately tackle webcare. More specifically, our findings on which strategies benefit or hurt business performance can help managers generate adequate prompts to be used by these language models for the task of replying to reviews.

6. Limitations and further research
Although the sample used in this study was reasonably diverse in terms of hotel type and size, the study should be replicated with a larger sample of hotels from a broader geographic area to enhance its external validity. Replicating it in a cross-cultural context with other tourism industries would test the findings’ generalizability.

Another area for further research is to explore which webcare strategies work best for which types of reviews, i.e. investigate how review characteristics moderate the effects of different webcare strategies. For instance, Allard et al. (2020) find that consumers are empathetic towards firms that receive unfair reviews, showing that review readers are also affected by the arguments used in the reviews. Future studies should also investigate the effects of combining multiple strategies on business performance, since in practice responses often use more than one strategy. Previous research points out that using several strategies in one response might yield the most positive results on consumers (van Hooijdonk and Liebrecht, 2021; Dens et al., 2015) and, consequently, affect future business. More research is needed to understand how tone of voice influences business performance. For instance, previous research finds that consumers appreciate interactions containing humor or comedy attempts from brands (Warren et al., 2018). Tone effects may be moderated further by the hotel brand, e.g. emojis may be effective for a hotel perceived as “fun”, but not for a business hotel perceived as “serious”. Future studies could study the effects of timeliness and tailoring on business performance since due to limitations in the data these strategies could not be tested.

Notes
1. One could question whether it would be better to control for observable causal factors such as the hotel’s average review rating over the past year (valence). It turns out that a hotel’s valence is mostly stable over time and aliased with the hotel_id, e.g. hotel 3 consistently has the highest star rating and
hotel 1 consistently has the worst. The hotel dummies explain 94.9% of the variation in valence, and a model with both has generalized VIF = 39.9. Hotel dummies are preferred over valence because they capture unobserved factors that vary between hotels.

References


Appendix
The supplementary material for this article can be found online.

Corresponding author
Ana Isabel Lopes can be contacted at: a.i.loureiro.lopes@vu.nl

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com