How fakes make it through: the role of review features versus consumer characteristics

Shabnam Azimi
Department of Marketing, Loyola University Chicago, Chicago, Illinois, USA
Kwong Chan
Department of Marketing, Northeastern University, Boston, Massachusetts, USA, and
Alexander Krasnikov
Department of Marketing, Loyola University Chicago, Chicago, Illinois, USA

Abstract

Purpose – This study aims to examine how characteristics of an online review and a consumer reading the review influence the probability that the consumer will assess the review as authentic (real) or inauthentic (fake). This study further examines the specific factors that increase or decrease a consumer's ability to detect a review's authenticity and reasons a consumer makes these authenticity assessments.

Design/methodology/approach – Hypothesized relationships were tested using an online experiment of over 400 respondents who collectively provided 3,224 authenticity assessments along with 3,181 written self-report reasons for assessing a review as authentic or inauthentic.

Findings – The findings indicate that specific combinations of factors including review valence, length, readability, type of content and consumer personality traits and demographics lead to systematic bias in assessing review authenticity. Using qualitative analysis, this paper provided further insight into why consumers are deceived.

Research limitations/implications – This research showed there are important differences in the way the authenticity assessment process works for positive versus negative reviews and identified factors that can make a fake review hard to spot or a real review hard to believe.

Practical implications – This research has implications for both consumers and businesses by emphasizing areas of vulnerability for fake information and providing guidance for how to design review systems for improved veracity.

Originality/value – This research is one of the few works that explicates how people assess information authenticity and their consequent assessment accuracy in the context of online reviews.

Keywords Online review, Fake review, Deception detection, Authenticity, Negative review

Paper type Research paper

> It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.
>
> Josh Billings Encyclopedia and Proverbial Philosophy of Wit and Humor (1874).

Introduction

Consumers frequently use online reviews as a key input into their decision-making (Cheung *et al.*, 2012; Kannan and Li, 2017). However, there is mainstream recognition in the media that inauthentic online reviews are not only pervasive (Ott *et al.*, 2012), but arguably inevitable because of the presence of clear economic incentives (Luca and Zervas, 2016). Given the importance of online reviews in consumer decisions, it is important for consumers to be able to distinguish real reviews from fake ones. Yet, past research indicates that human judges are notoriously bad at recognizing fake online reviews (Kim *et al.*, 2015; Yoo and Gretzel, 2009) with correct identification rates barely different from chance (Bond and Depaulo, 2006).

In the ongoing effort to provide readers of online reviews with more truthful information, research in communication, information systems, computer science and linguistics have used linguistic signals to train algorithms and detect deception (Gokhman *et al.*, 2012; Grazioli and Wang, 2001; Hancock *et al.*, 2008; Jindal and Liu, 2007; Mukherjee *et al.*, 2012; Ott *et al.*, 2011; Xiao and Benbasat, 2011). Such algorithms have been shown to identify fake reviews substantially better than humans. Nevertheless, advancements in these detection methods have largely failed to address the question of why readers are unable to systematically detect deception in reviews.

While most prior research examines linguistic signals used by writers of fake reviews, we emphasize readers rather than writers and the mechanism behind how readers make authenticity assessments. Relatively few studies have focused on testing the verbal cues relied upon by perceivers/readers when making authenticity judgments (Larrimore *et al.*, 2011; Toma and D'Angelo, 2015), and there is no clear consensus

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/0736-3761.htm**

Journal of Consumer Marketing
39/5 (2022) 523–537

[DOI 10.1108/JCM-04-2021-4597]

Received 8 April 2021
Revised 31 December 2021
31 March 2022
Accepted 24 June 2022

regarding the role of human perceiver characteristics in detecting deception. This is despite the fact people play a key role in filtering reviews before they are even seen online. In a world increasingly defined by automated algorithms, human judges still fulfill the necessary first step in labeling content as fake or real. These labeled data are in turn used to train machine learning algorithms (Wells and Alpert, 2018). Thus, without an understanding of how people are fooled, we are left with use of automated techniques that are largely post hoc in their reliance on historically trained data to filter new incoming information.

Drawing upon truth-default theory (TDT) (Levine, 2014) and the elaboration likelihood model (Petty and Cacioppo, 1986), this research aims to help build a cohesive image of how review features and consumer characteristics collectively predict consumer assessments of review authenticity. Using an established gold standard hotel review data set (Ott *et al.*, 2011) and an online experiment, we collected responses from 403 participants who provided 3,224 authenticity assessments. These same respondents provided 3,181 written self-report reasons for assessing a review as authentic (real) or inauthentic (fake). This design allows us to simultaneously examine how review-related variables, such as valence, length and readability, content type (affective or cognitive) and consumer-related variables, such as personality and demographics, shape assessments of review authenticity. We then determine what leads a person to be more likely to *correctly* assess a review's authenticity. Our findings highlight important differences in the way this process works for information that is negative versus positive, consequently providing new insight into the consumer evaluation process. We identify areas of vulnerability for consumers and businesses through identifying factors that can make a fake review hard to spot or a real review hard to believe.

Conceptual framework and hypotheses

Consumers are most influenced by information they deem to be trustworthy (Cheung *et al.*, 2008) and a precondition for trustworthiness is that reviews are judged to be real (Reimer and Benkenstein, 2016), credible (Cheung *et al.*, 2012), believable (Wathen and Burkell, 2002), authentic (Majumdar and Bose, 2018) and accurate (Filieri and McLeay, 2013). In this research, we use the term authenticity assessment to refer to this judgment process. In the following sections, we first develop hypotheses related to how people exhibit general bias in assessment of authenticity by adapting a well-known theory of deception detection, TDT; we then present our hypotheses related to review-related variables associated with authenticity assessment. We further review the literature related to consumer-related variables associated with authenticity assessment and explain that formulating hypotheses is made difficult because of a lack of prior consensus. We instead, pose a research question to assess consumer-related effects. The relationships presented are summarized in Figure 1.

Truth-default theory

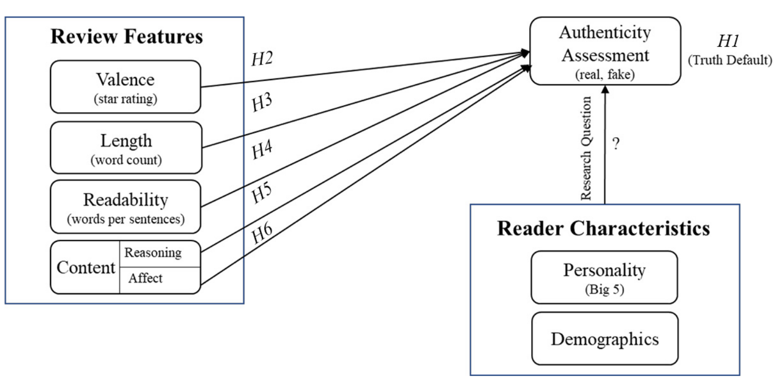
People are known to assume others tell the truth independent of actual honesty, a classic phenomenon that is called truth bias (Zuckerman *et al.*, 1981). Drawing upon truth bias, Levine's (2014) TDT describes how people make truth or lie judgments. TDT posits that people passively assume others are honest in their communication unless suspicion concerning honestly is triggered, whereupon bias shifts away from truth. However, even under conditions of suspicion, people remain fundamentally truth biased. The presumption of honesty disproportionately affects accuracy in judgment of truths versus judgment of lies, such that truths are often correctly identified as honest but lies are often also misidentified as truth (Levine *et al.*, 1999). Therefore, the truth–lie base rate or "the relative prevalence of deception and honesty in some defined environment" is a key variable that affects human judgment accuracy (Levine, 2014).

Deceptive reviews are highly prevalent and are often not discernible from genuine comments (Hancock *et al.*, 2008). Thus, even if consumers are aware of fake online reviews, based on the core premise in TDT and past research suggesting truth bias in the online review context (Kronrod *et al.*, 2017), we still expect bias toward judging reviews as real:

H1. People are more likely to assess reviews as real.

If consumers are biased toward truth judgments, specific reasons for this bias may be detectable and relatable to the truth-default mode. The purpose of subsequent sections is to assess which review features and consumer characteristics affect consumer truth judgments.

Figure 1 Conceptual model and hypotheses



Impact of review features on assessments of authenticity

We adopt Petty and Cacioppo's (1986) elaboration likelihood model (ELM) to identify which review features affect consumer perceptions of review authenticity. ELM is extensively used in e-WOM literature in understanding how consumers perceive online reviews. ELM suggests two processing routes are involved in persuasion: the central route, which entails a high level of cognitive processing and more effort to evaluate information, and the peripheral route, which involves less cognitive effort and use of simple short cuts to assess information. The use of central or peripheral cues depends on individual ability, willingness and motivation (e.g. involvement) to process information. Past research suggests that evaluation of believability of online reviews can be done based on both central and peripheral routes (Cheung *et al.*, 2012; Filieri and McLeay, 2013). Review attributes such as quantitative star ratings are processed through a peripheral route (Filieri and McLeay, 2013), whereas information contained in review text is processed through a central route (Agnihotri and Bhattacharya, 2016). The focus of the current study is on review valence (i.e. star rating) and length as representations of peripheral cues and review readability and affective versus cognitive types of content as representations of central cues in consumer assessments.

Review valence

Past research suggests that when consumers have prior positive brand attitudes, they find positive online reviews more persuasive than negative ones, as such when reviews confirm their original beliefs (Mafael *et al.*, 2016). However, irrespective of brand attitude, the extant literature provides support for a type of negativity bias in the online context where negative word of mouth is more influential than positive word of mouth (Casaló *et al.*, 2015; Chevalier and Mayzlin, 2006; Pan and Chiou, 2011; Park and Lee, 2009). As there are a larger number of positive reviews than negative reviews online, readers generally assume negative reviews are more informative (Ba and Pavlou, 2002; Pavlou and Dimoka, 2006). In addition, according to Chen and Lurie (2013), positive reviews are less likely to be based on actual experience and more likely to be written for self-serving purposes such as trying to confirm and feel good about one's decision. As such, readers may detect such bias in positive reviews which make these reviews seem less reliable. We therefore hypothesize:

H2. Negative reviews are more likely than positive reviews to be assessed as real.

Review length

Larrimore *et al.* (2011) found that use of extended narratives in loan descriptions can increase trust and improve fund raising success. Description length has been found relevant in other domains including an association with higher selling prices on eBay (Flanagin, 2007) and a perception that restaurant reviews are more useful (Cheng and Ho, 2015). These findings provide support for the positive impact of message length on reader perception of authenticity. Consequently, we expect:

H3. Longer online reviews are more likely to be assessed as real.

Readability

Readability refers to information understandability (Banerjee *et al.*, 2015), ease of language interpretation (Wang and Strong, 1996) and clarity of speech (Toma and Hancock, 2012). Research suggests that increased readability of an online review is positively related to its adoptability (Filieri and McLeay, 2013) and credibility (Ghose and Ipeirotis, 2011). Readers of online reviews may therefore better comprehend and interpret information from reviews which are easier to read. Previous research has operationalized readability using multiple indices including words per sentence (Ngo-Ye *et al.*, 2016) and suggests that lower readability reduces the perception of review authenticity (Toma and Hancock, 2012):

H4. Reviews that are more readable are more likely to be assessed as real.

Cognitive and emotional review content

Type of content is an important central cue that can impact consumer judgments of authenticity. We focus on the roles of cognitive and affective words on consumer evaluation of a review. Cognitive content includes a greater number of causation words such as "because," "effect" and "hence" which add specificity and detail to a review (Hancock *et al.*, 2008). Prior work shows that such reasoning reduces decision uncertainty (Berger and Calabrese, 1975) and can have a positive effect on perceptions of trustworthiness (Cheung *et al.*, 2008; Reimer and Benkenstein, 2016). Moreover, causation terms may be perceived as authenticity indicators because they add more specificity and details in the reviews (Hancock *et al.*, 2008). We therefore expect causation words to increase a reader's perception a review is authentic:

H5. Reviews that use more causation terms are more likely to be assessed as real.

Felbermayr and Nanopoulos's (2016) study highlighted the importance of emotional content on the perceived helpfulness of online reviews and indicated that trust, joy and anticipation are the most influential emotional dimensions. Zablocki *et al.* (2019) examined the impact of emotions on attitude and found a positive effect for positive emotional content and a negative effect for negative or mixed emotional content. The impact of affect words on credibility perception has received limited attention; however, there is some evidence indicating that increased use of affect words (e.g. "sweet," "nice," "ugly," "hurt") is attributed to exaggeration of review sentiment and can lower credibility (Ott *et al.*, 2013). Hence, we predict:

H6. Reviews that use more affect words are less likely to be assessed as real.

Impact of consumer characteristics on assessments of authenticity

In addition to message-related variables, the literature concerning communication persuasiveness and informational influence highlight the importance of recipient-related variables (Holvand, 1959; Petty and Cacioppo, 1981). Despite this importance, there is scant research examining the roles that personality types and demographics play in assessing information accuracy and identification of deception in online reviews. Regarding the role of age, Munzel (2015) found that younger groups (with an average age of 23 years) perceive online restaurant reviews to be more trustworthy than older groups (average age of 37).

Personality types can be characterized by the “Big-Five” OCEAN factors (John and Srivastava, 1999) that consist of openness, conscientiousness, extraversion, agreeableness and neuroticism. Amongst the five personality factors, past research only provides evidence for the effect of openness, agreeableness and neuroticism on reader perceptions of accuracy of online reviews. In one of the few studies in this area, Enos *et al.* (2006) found that the neuroticism and agreeableness of human judges were positively correlated with frequency of guessing statements as truthful, whereas openness was positively correlated with frequency of guessing statements as deceptive.

Because of a lack of conclusive evidence regarding the role of reader characteristics, we pose a research question to address the current deficit in research concerning the role receiver personality types and demographics may play in the assessment of deception in online reviews:

RQ1. Do consumer personality types and/or demographics affect assessments of review authenticity?

Methodology

Data set of reviews

To measure a subject’s ability to assess review authenticity, we used the gold standard opinion spam data set of 1,600 reviews for 20 hotels in the downtown area of Chicago (Ott *et al.*, 2011; Ott *et al.*, 2013). This database provides a balanced set of reviews for each hotel by positive/negative valence and fake/real veracity. Ott *et al.* (2011) extracted real reviews from Expedia, Hotels.com, Orbitz, Priceline, Tripadvisor and Yelp. Fake reviews were gathered using the Amazon Mechanical Turk (AMT) crowdsourcing website. To gather fake reviews, the researchers in Ott *et al.* presented AMT workers with the name and website of one of the 20 hotels and asked them to assume they work for the hotel’s marketing department. They were then instructed to either “write a fake positive review as if they were a customer to be posted on a travel review website” or “write a fake negative review of a competitor’s hotel to be posted online.” The review needed to sound truthful.

Experimental design

Use of the gold standard data set by Ott *et al.* (2011, 2013) enabled us to manipulate review valence and veracity. The reviews in the study are a random selection from the Ott *et al.* (2011, 2013) gold standard review data set. Multiple hotels are represented in the study design to avoid systematic variance stemming from hotel-related characteristics. For each of 20 hotels we randomly compiled eight reviews balanced across valence and veracity by sampling two negative-fake, two negative-real, two positive-fake and two positive-real reviews per hotel from the mentioned data set. A sample review from each of the four categories is presented in the Appendix 1. In total, 403 participants currently residing in the USA (57.2% female; age 17–24: 24%, 25–34: 41%, 35–44: 17%, 45–54: 8% and >55: 9%) were recruited from the AMT website. Participants were asked to imagine they are going to visit Chicago and needed to choose a hotel for their stay. Each subject was randomly assigned to read eight reviews on one of the 20 Chicago hotels and guess whether each review was fake or real and provide a reason along with their guess. The reviews are presented in random order to respondents to eliminate order effects. They were then asked to answer questions related to OCEAN personality dimensions, demographics and prior familiarity with Chicago hotels. We designed the study so that each review was read and judged by 20 subjects.

Independent and dependent variables

Review veracity is the first independent variable which is based on the originally fake or real reviews taken from the data set. To manipulate review valence, star ratings of one or two were presented along with negative reviews and star ratings of four or five were presented along with positive reviews. The approach we apply to manipulate our independent variables is consistent with past research (Sen and Lerman, 2007; Zablocki *et al.*, 2019; Mafael *et al.*, 2016). A subject’s assessment of whether a review is fake or real was our dependent variable which was measured using a categorical variable with 0 = fake and 1 = real. A subject’s fake/real identification accuracy was calculated using a binary scale of 0 = inaccurate identification and 1 = accurate identification based on the match between the originally labeled veracity of the review and reader assessment of review veracity.

To extract the linguistic cues used in our sample reviews, we used the commonly applied (Larrimore *et al.*, 2011) Linguistic Inquiry and Word Count (LIWC) 2015 software. LIWC is a text analysis application that analyzes emotional, cognitive and structural components of verbal and written samples (Pennebaker *et al.*, 2015). Applicability of LIWC has been widely shown in computer science, marketing, psychology and communication and it has been successfully used to predict many psychological outcomes such as writers’ personality, deception and helpfulness (Fast *et al.*, 2016; Hancock *et al.*, 2008; Larrimore *et al.*, 2011; Li and Chignell, 2010; Pennebaker and King, 1999; Toma and Hancock, 2012; Zhu *et al.*, 2020).

Mean word count, average words per sentence and mean percentages of causation and affect-words across the four categories of negative-fake, negative-real, positive-fake and positive-real reviews are reported in Table 1.

A subset of the 44-item Big Five Inventory was used (John *et al.*, 1991) to measure personality dimensions. All five personality constructs were measured on a five-point scale and had Cronbach’s alphas greater than 0.8 (Appendix 2). We also measured prior familiarity with Chicago hotels as a control variable in our study. We found 64% of the subjects had never heard of the hotel and 36% either had heard of or had stayed at the hotel for which they were reading reviews.

Data structure and modeling approach

Our experimental design implies repeated observations of the same subjects and captures binary outcomes (i.e. fake versus real). Standard logistic modeling may not be effective because of clustering (within subject and hotel chains) and overdispersion (Hu *et al.*, 1998). As such, we used the generalized estimation equations (GEE) (Lalonde *et al.*, 2013)

Table 1 Mean of linguistic variables for the four conditions

Linguistic variables	Negative fake *M*	Negative fake SD	Negative real *M*	Negative real SD	Positive fake *M*	Positive fake SD	Positive real *M*	Positive real SD
Word count	174.89	75.07	197.55	106.47	117.86	62.68	125.07	69.27
Words per sentence	17.92	3.43	16.84	7.11	16.15	4.02	16.30	13.23
Causation (because, effect)	1.09	0.75	1.07	0.81	0.51	0.83	0.65	0.77
Affect words (love, happy, cry, annoyed, kill, sweet)	4.44	1.80	4.88	3.03	7.49	2.76	7.51	2.97

method to model the outcome variable: π_{ijk} probability that subject i classifies review j about hotel chain k as real [equation (1)]:

$$Log\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = \alpha_0 + \alpha_1 * WC_{jk} + \alpha_2 * WPS_{jk} + \alpha_3 * Causation_{jk} + \alpha_4 * Affect\ words_{jk} + \alpha_5 * Neuroticism_i + \alpha_6 * Extraversion_i + \alpha_7 * Agreeableness_i + \alpha_8 * Openness_i + \alpha_9 * Conscientiousness_i + \alpha_{10} * Gender_i + \alpha_{11} * Age_i + \alpha_{12} * Education_i + \alpha_{13} * Prior\_Familiarity_{ik} + \varepsilon_{ijk} \quad (1)$$

where:

π	= is the probability that subject i identifies review j of hotel chain k as real (1) versus fake (0),
WC	= is the word count of review j of hotel chain k,
WPS	= is the words per sentence,
Causation	= is the number of causation terms,
Affect Words	= is the number of affect words,
Neuroticism, Extraversion, Agreeableness, Openness, Conscientiousness	= is the score of Big Five's dimensions,
Gender	= is the respondent's sex,
Age	= is the age of respondent,
Education	= is the education level; and
Prior Familiarity	= is the familiarity of subject i with hotel chain k.

The GEE parameter estimates in equation (1) are derived using the correlation structure $R(\gamma)$ between predictors, where γ is vector of correlation parameters that fully captures correlations among variables (for brevity we omitted subscripts for hotel chain and review). When fitting GEE model in equation (1) and accounting for autocorrelation between responses, the marginal response variance V_i for subject i may be estimated as:

$$V_n(\gamma, \phi) = \phi A_i^{1/2} R_i(\gamma) A_i^{1/2} \quad (2)$$

where A_i is a diagonal matrix representing the variance under the assumption of independence and ϕ is the overdispersion factor. The GEE for I subjects in model (1) will have the following form:

$$U(\alpha) = \sum_{i=1}^{I}\left(\frac{\partial \pi_i(\alpha)}{\partial \alpha}\right)^T V_i^{-1}\left(\hat{\gamma}(\alpha),\ \hat{\phi}(\alpha)\right)\{Real_i - \pi_i(\alpha)\} = 0 \quad (3)$$

where Real equals 1 (0) if a subject identifies review as real (fake).

Solving these equations using dispersion parameter ϕ and correlations γ provides estimates for equation (1). Moreover, the efficiency of the GEE estimates may be enhanced if a working correlation matrix is selected using quasi-likelihood criteria (Pan, 2001). We centered linguistic predictors at their grand means and Big Five and subject characteristics by each review (Enders and Tofighi, 2007) to model the heterogeneity of subjects participating in the experiment.

Qualitative analysis

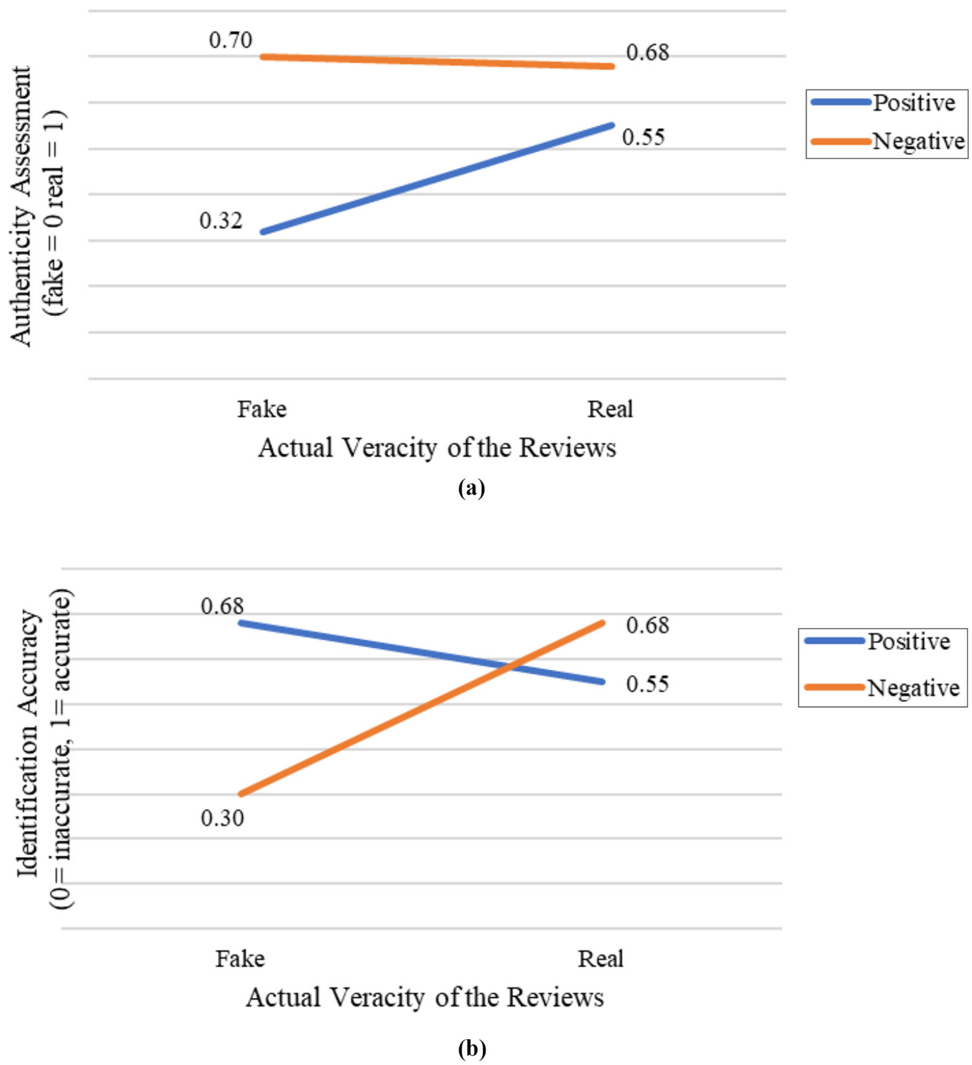
In addition to indicating if they thought a review was fake or real, respondents were also asked to provide a rationale for each guess in an open-ended section for each review that was assessed. Respondents provided 3,181 (out of a possible 3,224) text passages. Two authors iteratively compared these passages and designed a preliminary coding scheme whereby each passage could be classified along multiple reasons the reader thought a review was fake or real. Using this scheme, two research assistants then independently coded each of the passages and also were allowed to create new coding classifications when deemed necessary. Discrepancies in coding choices between the two research assistants were resolved by having a third research assistant independently code the passages with discrepancy classifications and the majority vote taken. A fourth research assistant aggregated the data in NVivo, assigning case roles to units of analysis according to people (respondents), and reviewed attributes including fake/real authenticity states. Nodes were assigned to coded values derived from the typed passages provided by respondents. Qualitative codes such as "too positive," "facts seem fake" and "reads like an advertisement" are examples of these nodes.

Results

Average consumer authenticity assessment

The average believability for each of the four types of reviews is shown in Figure 2(a). Overall, the mean reader authenticity assessment of a review as either fake or real across 3,224 observations indicated a bias toward real that differed significantly from 50/50 split ($M = 0.56$, $t = 6.81$, $p < 0.01$, where 0 = fake and 1 = real). This indicates readers have a greater tendency to assess reviews as real and supports *H1*. *H2* posits people evaluate negative reviews as more authentic than positive reviews. Our results support *H2* as the average authenticity assessment for negative reviews is significantly

Figure 2 Review valence and review veracity effects on authenticity assessment and identification accuracy



Notes: (a) Authenticity assessment means; (b) Identification accuracy means

higher than authenticity assessment for positive reviews (M = 0.69 versus M = 0.43, t = 14.87, $p < 0.05$). Interestingly, the most believable reviews are negative fake (M = 0.70) or negative real (M = 0.68) and the least believable reviews are positive fake (M = 0.32).

Effects of review and consumer characteristics on authenticity assessment

The relationship between review or consumer characteristics and the dependent variable are examined through the GEE approach. We ran this model for the entire sample (Model A). For completeness, we also show subsample results for positive review assessments (Model B) and negative review assessments (Model C) to determine if similar effects exist within review valence types. The parameter estimates in our GEE model are presented as log odds ratios for our binary outcomes. Parameter estimates, Z scores and the Quasilikelihood under the Independence model Criterion (QIC) goodness of fit statistic related to these three models are shown in Table 2.

Consistent with theory and unless otherwise hypothesized, we ran one-tailed tests to examine the hypothesized relationships. In Model A, results were congruent with predictions in *H3* and *H5*, higher incidence of word count (β = 0.028, $p < 0.05$) and causation terms (β = 0.045, $p < 0.01$) increase the likelihood a review is assessed as real by a reader. In contrast, greater use of affect words (β = –0.117, $p < 0.01$) increases the chances a review is assessed as fake, consistent with *H6*. However, *H4* is not supported as words per sentence did not impact reader assessments of whether a review was fake or real (β = –0.07, $p > 0.10$).

We further explored the effects of linguistic variables for positive (Model B) and negative reviews (Model C). Surprisingly, we found different effects for most linguistic cues on reader assessments of positive versus negative reviews. Word count (β = –0.029, $p < 0.10$) and affect (β = –0.040, $p < 0.10$) only impacted reader assessment of negative reviews, whereas words per sentence had opposite effects on assessment of

Table 2 Effects of review and consumer characteristics on consumer assessment

Variable	Expected effect	Model A Entire sample B (Z)	Model B Positive reviews B (Z)	Model C Negative reviews B (Z)
Intercept		–0.546 (9.39)***	–0.845 (7.99)***	–0.299 (3.84)***
Review characteristics				
Word count	*H3* (+)	0.028 (1.95)**	0.040 (1.08)	–0.029 (1.77)*
Words per sentence	*H4* (–)	–0.007 (0.45)	–0.069 (1.77)*	0.038 (1.73)*
Causation	*H5* (+)	0.045 (3.04)***	–0.004 (0.15)	–0.005 (0.26)
Affect	*H6* (–)	–0.117 (6.15)***	0.007 (0.23)	–0.040 (1.87)*
Consumer characteristics	RQ			
Big Five personality				
Neuroticism		0.025 (1.34)	–0.002 (0.07)	0.034 (1.49)
Extraversion		–0.001 (0.06)	0.004 (0.13)	–0.005 (0.23)
Agreeableness		0.035 (1.96)**	0.077 (2.15)**	0.013 (0.58)
Openness		0.001(0.07)	–0.009 (0.29)	0.007 (0.33)
Conscientiousness		0.004 (0.23)	–0.006 (0.20)	0.009 (0.42)
Demographics				
Gender		0.014 (0.43)	0.030 (0.50)	0.008 (0.19)
Age		–0.028 (2.06)**	–0.017 (0.74)	–0.039 (2.14)**
Education		–0.005 (0.34)	–0.021 (0.79)	0.006 (0.29)
Control variable				
Prior familiarity		0.006 (0.40)	–0.007 (0.22)	0.015 (0.71)
QIC		3237.376	1623.681	1632.390

Notes: Dependent variable: consumer assessment (fake = 0, real = 1); gender (male = 0, female = 1); *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

positive (β = –0.069, $p < 0.10$) versus negative reviews (β = 0.038, $p < 0.10$). Overall, these results suggest that positive reviews with shorter sentences (more readable) are perceived as more real, whereas negative reviews that are shorter, have longer sentences and use lower number of affect words are perceived as being more real.

Results related to reader characteristics (research question) showed that irrespective of review valence, agreeableness increases the chances a review will be assessed as real (β = 0.035, $p < 0.05$). However, Models B and C suggest that this positive assessment only works in favor of positive reviews (β = 0.077, $p < 0.05$). We also found that overall, older people are less likely to judge reviews as real (β = –0.028, $p < 0.05$) compared to younger people. Models B and C further suggest that older adults are more skeptical about the authenticity of negative reviews (β = –0.039, $p < 0.05$) but their assessments of positive reviews are not significantly different from younger adults (β = –0.017, $p > 0.10$). No main effects of other personality-related or demographic variables were significant.

Does perception correspond to accuracy?

As Figure 2(b) shows, average accuracy in identification of real reviews (M = 0.61) was significantly higher than average accuracy in identification of fake reviews (M = 0.49, t = 6.85, $p < 0.05$). Our results further demonstrate that readers were more successful in identifying positive reviews (M = 0.62) than negative ones (M = 0.49, t = 7.42, $p < 0.05$). We compared identification accuracy of positive with negative reviews within each of the two categories of fake and real review valence. When the reviews were real, negative reviews (M = 0.68) were more accurately identified than positive reviews (M = 0.55, t = 5.36, $p < 0.05$). For fake reviews, readers had higher average accuracy in identifying positive (M = 0.68) than negative reviews (M = 0.30, t = 15.26, $p < 0.05$). These results provide insight into factors that affect whether a review is viewed to be authentic. As a post hoc analysis, we next examine how correct these judgments of authenticity are and specifically whether review features or consumer characteristics are related to identification accuracy. To tease out the role of review valence, we tested these effects for samples of positive and negative reviews as well as the entire sample. The results are shown in Table 3.

The results show that increased use of affect words improved accuracy irrespective of review valence (β = 0.046, $p < 0.001$). However, the effects of other review characteristics were conditional upon review valence. Readers had higher accuracy in identifying negative reviews that are longer (β = 0.092, $p < 0.001$) and use more causation terms (β = 0.049, $p < 0.10$) and positive reviews which had shorter sentences (β = –0.045, $p < 0.05$) and used less causation terms (β = –0.037, $p < 0.10$).

Results in Models 1–3 also show that readers with different characteristics have different levels of accuracy in identifying fake or real reviews. Readers with higher extraversion scores were overall less accurate (β = –0.031, $p < 0.10$) and readers with higher openness scores were more accurate (β = 0.035, $p < 0.05$). However, estimates for Models 2–3 indicate that these effects were only significant in identifying positive reviews with no better/worse accuracy for negative reviews. As might be expected because of knowledge endowment effects, our control variable, prior familiarity with Chicago hotels, had a positive effect on accuracy in identifying positive reviews with no effect for negative reviews. Other factors including personality traits, gender and education had

Table 3 Effects of review and consumer characteristics on identification accuracy

Variable	Model 1 Entire sample	Model 2 Positive reviews	Model 3 Negative reviews
	B (*Z*)	*B* (*Z*)	*B* (*Z*)
Intercept	–0.618 (10.35)***	–0.512 (6.28)***	–0.792 (8.64)***
Review characteristics			
Word count	0.016 (0.96)	–0.039 (1.31)	0.092 (4.37)***
Words per sentence	–0.027 (1.37)	–0.045 (1.94)**	0.049 (1.49)
Causation	–0.025 (1.46)	–0.037 (1.66)*	0.049 (1.78)*
Affect	0.046 (2.73)***	–0.025 (1.17)	0.047 (1.37)
Consumer characteristics			
Big Five personality			
Neuroticism	–0.004 (0.19)	–0.002 (0.08)	–0.009 (0.32)
Extraversion	–0.031 (1.79)*	–0.038 (1.69)*	–0.025 (0.92)
Agreeableness	0.001 (0.02)	–0.016 (0.71)	0.023 (0.82)
Openness	0.035(1.92)**	0.056 (2.38)***	0.004 (0.15)
Conscientiousness	–0.003 (0.17)	–0.014 (0.60)	0.014 (0.50)
Demographics			
Gender	0.035 (1.05)	0.044 (0.94)	0.021 (0.42)
Age	–0.014 (1.08)	–0.028 (1.55)	0.004 (0.22)
Education	–0.001 (0.09)	–0.007 (0.34)	0.006 (0.26)
Control variable			
Prior familiarity	0.034 (2.22)**	0.049 (2.57)***	0.011 (0.49)
QIC	3236.625	1627.533	1621.843

Notes: Dependent variable: identification accuracy (not accurate= 0, accurate = 1); gender (male = 0, female = 1); $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

no impact on reader ability to accurately assess the authenticity of positive or negative hotel reviews.

For robustness checks, we assessed whether a respondent's authenticity assessment and identification accuracy were affected by hotel attributes mentioned in reviews. We coded all hotel reviews based on the most common themes mentioned in reviews such as room cleanliness, friendliness of staff, speed of service, location and conditions of hotel amenities. These attributes were in line with past research in tourism and hotel marketing (Kusumasondjaja *et al.*, 2012). For each of these themes, we defined a binary variable for whether a review features that theme or not. Next, we included five dummies in equation (1) and tested their impact on two outcomes (perception of whether a review is real and accuracy in guessing) for various samples (entire sample, positive reviews and negative reviews). The results suggest that the effects of these variables were not statistically significant and the results of alternative models including these hotel attributes were consistent with those reported in the Tables 2 and 3.

Qualitative results: why respondents assessed a review as fake or real

Our collection of self-reported qualitative reasons for why a respondent viewed a review as real or fake indicated that in many cases a person may have a strong belief that a fake review should exhibit certain characteristics when in fact these characteristics are irrelevant to review authenticity or even indicative of a real review. One such example is the reason: "I don't think a fake review would be negative." In fact, this type of justification is not rare and the belief that fake reviews are likely to be positive was provided by several respondents in various forms including: "Too much detail to be fake. And why would someone want to give a fake negative review?" and "Why would anyone create a fake review and give it 1 star. People that have bad experiences don't need to fake a review."

There are many criteria used by people in assessing authenticity that lead to inaccurate judgments. Illustrative examples of such reasons that reflect the most and least effective rationales for assessment of a review as real are shown in Figure 3. The upper line with higher data values indicates the number of times the reason was given for assessing a review as "real" and the lower line with lower value data points indicates the number of times the reason was correctly associated with an actual real review (i.e. assessed as real when the review was truly real). "Genuine tone" was mentioned in 181 guesses that a review was real, however this reason was only correct in 110 instances (as an incidence baseline, recall that we have comments linked directly to 3,181 guesses). The large gap between lines for reasons such as "Genuine tone" and "Specific details provided" indicate these reasons are unlikely to be reliable criteria for evaluating a review as real, whereas closer lines such as "Has negative aspects" and "Has both positive and negative aspects" are valid reasons for evaluating a real review as real, because these latter reasons are more likely to lead to a correct assessment that a review is real.

An illustrative sample of the most ineffective reasons to guess fake compared to the most effective is shown in Figure 4. "Fact seems fake" was mentioned 155 times as a reason for assessing a review as fake but this reason was only correctly associated with a fake guess 99 times. The larger distance between lines

Figure 3 Representative reasons for assessing a review as real

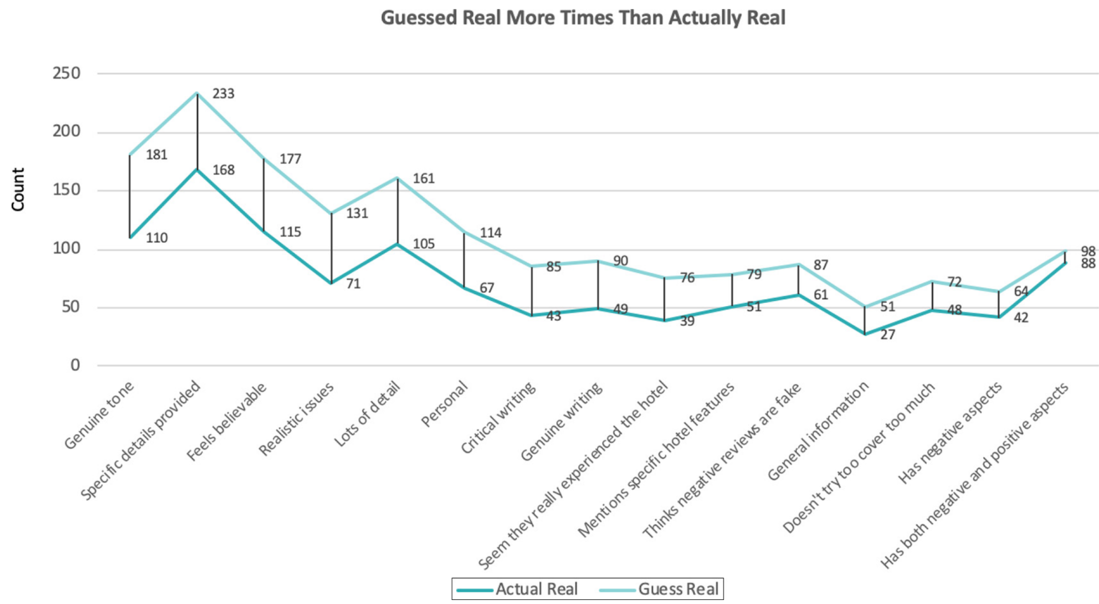
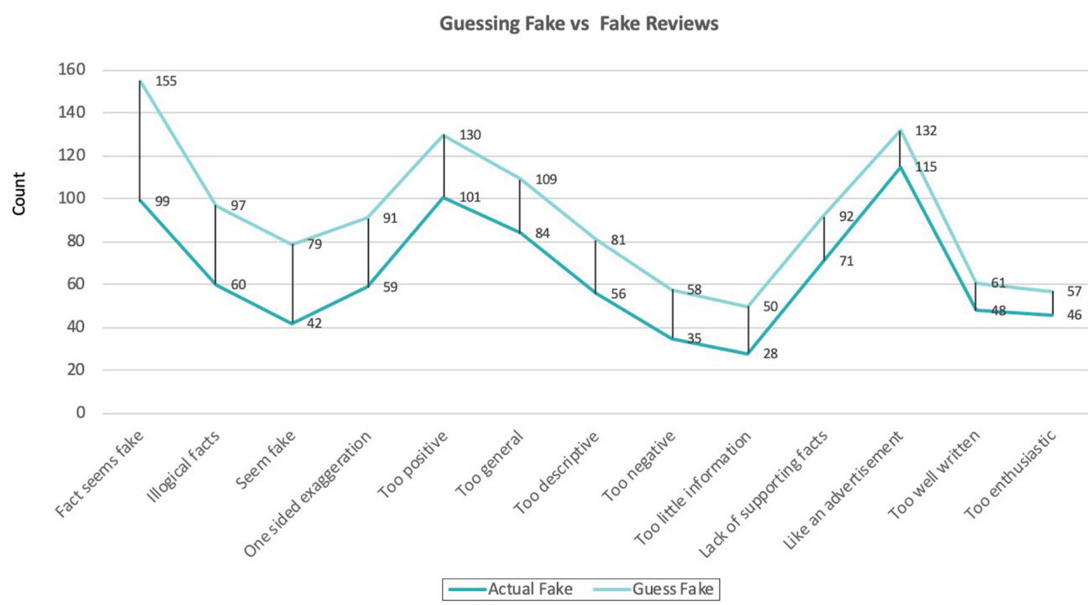


Figure 4 Representative reasons for assessing a review as fake



indicates reasons such as "Fact seems fake" and "illogical facts" are the least reliable justifications to evaluate a fake review as fake, whereas the smaller distance between lines indicates "Too well written" and "Too enthusiastic" are the most valid reasons to evaluate a fake review as fake.

Discussion

Studies across marketing, psychology, computer science and communications have found human judges are remarkably poor detectors of fake information. However, the degree to which this lack of capability stems from linguistic versus reader-centric factors has not been examined. This research is one of the few works that explicates the complicated process behind how people assess information authenticity and their consequent authenticity assessment accuracy in the context of online reviews. By examining the combined impact of review textual features and human reader characteristics, building upon major theories of TDT (Levine, 2014), the ELM (Petty and Cacioppo, 1986) and negativity bias (Kanouse and Hanson, 1972), we identified important differences in the way the authenticity assessment process works for negative versus positive information. One of our major contributions to the theory is unraveling the moderating role of valence in this process.

Our first finding, consistent with predictions of TDT (Levine, 2014) and research on negativity bias (Kanouse and Hanson, 1972; Sen and Lerman, 2007), is that people tend to assess reviews as real rather than fake and tend to view negative reviews as more authentic than positive reviews. When we measured how these biased perceptions translate into detection accuracy, we found significant differences between accuracy in identification of fake negative reviews (30% accuracy) and fake positive reviews (70% accuracy). Qualitative responses revealed a plethora of reasons for this difference including an expressed belief by individuals that there is little motivation to create a negative fake review. This reflects a naiveté in thinking that ignores the obvious justifications such as competitive hostility (Mayzlin *et al.*, 2014) and reviewer revenge. We find that not only do truth (*H1*) and negativity (*H2*) biases exist in judgment of online reviews but also that a combination of these effects will lead to the worst levels of accuracy when evaluating negative fake reviews. These results highlight the lesser-known potential impact of fake negative information on our everyday choices.

Why is a review assessed as real?

Our analysis identified different linguistic cues that drive authenticity assessments of positive versus negative information. Our findings can partly be explained by negativity bias as our readers were more critical in assessing the authenticity of positive reviews by trusting more readable examples, whereas they evaluated shorter negative reviews with longer sentences as more authentic. We further found that younger respondents found negative reviews to be more believable, whereas positive reviews are more believable for people who scored higher on agreeableness. Evidence of positivity bias in agreeable personalities is in line with past research (Augustine *et al.*, 2011).

Which personality types are better in identification of review?

We found that negativity bias influences people irrespective of their personality type. However, in analyzing judgment of positive reviews, we found those who score higher on extraversion have less success and those who score higher on openness score have more success in identification of reviews.

How do fakes make it through?

To tease out what factors lead to poor accuracy in identification of fake reviews, we did a detailed analysis of the effects of review and reader characteristics on identification accuracy across the four experimental conditions (see Appendix 3). The results (Models B and D) revealed that review-related features play a role in poor identification of negative fake reviews, whereas reader personalities impact poor identification of positive fake reviews. Shorter review length and decreased use of causation terms make a negative fake review hard to spot. In contrast, those individuals who are more agreeable and less open have a harder time spotting positive fake reviews.

What makes a real review hard to believe?

There are situations in which a review is genuine, but readers do not find it trustworthy. Models A and C in Appendix 3 indicate reader accuracy in identification of positive or negative real reviews. Negative effects indicate readers assessed those reviews as fake, when in fact they were real (hence they were real but not believable). We found that readability (as indicated by shorter sentences) has two opposite effects on believability of positive versus negative real reviews. Positive real reviews with longer sentences and negative real reviews with shorter sentences are less likely to be believed. In addition, negative real reviews with a smaller number of causation terms are less likely to be believed. We found no systematic effects in relation to personality dimensions and assessment of positive or negative reviews.

Managerial implications

This research has implications for both consumers and businesses by highlighting areas of vulnerability and providing guidance on how to use reviews more effectively. Through understanding human- and review-related factors that shape perceptions of review authenticity, consumers can learn their likely trigger points and businesses can learn about consumer perceptions. Businesses can learn from situations in which their real online reviews can be perceived as fake, or fake reviews taken to be real (possibility written by competitors). Consumers can be more informed about their misassessments of online reviews and therefore make better decisions.

Not only are fake negative reviews likely to make it through a human filter but they will also substantially influence their opinion (Basuroy *et al.*, 2003). Firm strategies to detect fake information and reviews are multifaceted. In addition to use of machine learning algorithms to automatically detect reviews of questionable authenticity, firms also hire human workers to mark fake content, run experiments on users and pursue legal strategies such as filing lawsuits against fake writers. Our results directly inform the first two methods by highlighting review features that increase the probability information is assessed as fake and illuminating which workers may be most able to detect false information. These are less risky than running live experiments on users that may attract negative publicity (BBC News, 2017) and costly lawsuits that must be filed across international jurisdictions and have uncertain outcomes.

Limitations and future research

As the first study to both examine consumer perceptions of online review authenticity and the process underlying accuracy in identifying fake reviews, this research has perhaps raised more questions than it has answered. The investigation of linguistic and personality cues has identified sources of bias from both signaling and information processing perspectives. However, we have yet to understand why certain linguistic cues stand out as significant determinants of accuracy identification nor do we know why their effects seem conditional on review sentiment. Our research may also have some methodological limitations. We captured major reviews or reader-related variables contributing to authenticity judgments through our conceptual model and qualitative section, but we do not exactly know how far readers attended and comprehended the reviews. Future research can study these effects in a more controlled experimental setting and explore use of more controlled text in manipulation of their independent variables. We are also aware of some shortcomings of the LIWC software. For instance, although words per sentence is commonly used in the literature as a measure of text readability (Ngo-Ye *et al.*, 2016;

Toma and Hancock, 2012), readability can be measured through other indices including Flesch grade level or Gunning fog index. In addition, LIWC's predefined dictionary misses certain topics (i.e. hotel attributes applicable to the current research) and those categories need to be identified and uploaded as specialized dictionaries (Taraban and Khaleel, 2019).

Past research shows that priming readers for existence of deception increases the likelihood of those reviews being judged as fake (Munzel, 2015), indicating that our results for reader assessments of the prevalence of fake reviews could be inflated. While several studies suggest fake hotel reviews are highly prevalent across platforms at up to 40% in some cases (Birchall, 2018), our study protocol had both a 50% prevalence of fake and 50% prevalence of negative reviews which is higher than most real-world scenarios. Further research may wish to choose fake and valence distributions that more accurately reflect real-world incidence. We also invite future research to extend our findings to other contexts such as fake news. It would be valuable to see how the skepticism toward news outlets may change the impact of fake information on consumer perceptions and judgments.

References

Agnihotri, A. and Bhattacharya, S. (2016), "Online review helpfulness: role of qualitative factors", *Psychology & Marketing*, Vol. 33 No. 11, pp. 1006-1017.

Augustine, A.A., Mehl, M.R. and Larsen, R.J. (2011), "A positivity bias in written and spoken English and its moderation by personality and gender", *Social Psychological and Personality Science*, Vol. 2 No. 5, pp. 508-515.

Ba, S. and Pavlou, P.A. (2002), "Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior", *MIS Quarterly*, Vol. 26 No. 3, pp. 243-268.

Banerjee, S., Chua, A.Y.K. and Kim, J.-J. (2015), "Using supervised learning to classify authentic and fake online reviews", *Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication*.

Basuroy, S., Chatterjee, S. and Ravid, S.A. (2003), "How critical are critical reviews? The box office effects of film critics, star power, and budgets", *Journal of Marketing*, Vol. 67 No. 4, pp. 105-117.

BBC News (2017), "Facebook's fake news experiment backfires", November 2017, available at: www.bbc.com/news/technology-41900877

Berger, C.R. and Calabrese, R.J. (1975), "Some exploration in initial interaction and beyond: toward a developmental theory of interpersonal communication", *Human Communication Research*, Vol. 1 No. 2, pp. 99-112.

Birchall, G. (2018), "TripAdvisor denies claims one in three reviews 'faked'", available at: www.news.com.au/technology/online/social/tripadvisor-denies-claims-one-in-three-reviews-faked/news-story/55243de188cc7f1fb2abb52fee3bac45 (accessed 12 June 2019).

Bond, C.F. and DePaulo, B.M. (2006), "Accuracy of deception judgments", *Personality and Social Psychology Review*, Vol. 10 No. 3, pp. 214-234.

Casaló, L.V., Flavián, C., Guinalíu, M. and Ekinci, Y. (2015), "Avoiding the dark side of positive online consumer reviews: enhancing reviews' usefulness for high risk-averse travelers", *Journal of Business Research*, Vol. 68 No. 9, pp. 1829-1835.

Chen, Z. and Lurie, N.H. (2013), "Temporal contiguity and negativity bias in the impact of online word of mouth", *Journal of Marketing Research*, Vol. 50 No. 4, pp. 463-476.

Cheng, Y.-H. and Ho, H.-Y. (2015), "Social influence's impact on reader perceptions of online reviews", *Journal of Business Research*, Vol. 68 No. 4, pp. 883-887.

Cheung, C.M., Lee, M.K. and Rabjohn, N. (2008), "The impact of electronic word-of-mouth: the adoption of online opinions in online customer communities", *Internet Research*, Vol. 18 No. 3, pp. 229-247.

Cheung, C.M.-Y., Sia, C. and Kuan, 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*, Vol. 13 No. 8, pp. 618-635.

Chevalier, J.A. and Mayzlin, D. (2006), "The effect of word of mouth on sales: online book reviews", *Journal of Marketing Research*, Vol. 43 No. 3, pp. 345-354.

Enders, C.K. and Tofighi, D. (2007), "Centering predictor variables in cross-sectional multilevel models: a new look at an old issue", *Psychological Methods*, Vol. 12 No. 2, pp. 121-138.

Enos, F., Benus, S., Cautin, R., Graciarena, M., Hirschberg, J. and Shriberg, E. (2006), "Personality factors in human deception detection: comparing human to machine performance", *Proceedings of the Annual Conference of the International Speech Communication Association*.

Fast, E., Chen, B. and Bernstein, M. (2016), "Empath: understanding topic signals in large-scale text", *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, May.

Felbermayr, A. and Nanopoulos, A. (2016), "The role of emotions for the perceived usefulness in online customer reviews", *Journal of Interactive Marketing*, Vol. 36, pp. 60-76.

Filieri, R. and McLeay, F. (2013), "E-WOM and accommodation: an analysis of the factors that influence travelers' adoption of information from online reviews", *Journal of Travel Research*, Vol. 53 No. 1, pp. 44-57.

Flanagin, A.J. (2007), "Commercial markets as communication markets: uncertainty reduction through mediated information exchange in online auctions", *New Media & Society*, Vol. 9 No. 3, pp. 401-423.

Ghose, A. and Ipeirotis, P.G. (2011), "Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23 No. 10, pp. 1498-1512.

Gokhman, S., Hancock, J.T., Prabhu, P., Ott, M. and Cardie, C. (2012), "In search of a gold standard in studies of deception", *Proceedings of the Workshop on Computational Approaches to Deception Detection*.

Grazioli, S. and Wang, A. (2001), "Looking without seeing: understanding unsophisticated consumers' success and failure to detect internet deception", *22nd international conference on information systems, New Orleans, LA*, pp. 193-203.

Hancock, J.T., Curry, L.E., Goorha, S. and Woodworth, M. (2008), "On lying and being lied to: a linguistic analysis of deception in computer mediated communication", *Discourse Processes*, Vol. 45 No. 1, pp. 1-23.

Holvand, C.L. (1959), "Reconciling conflicting results derived from experimental and survey results studies of attitude change", *American Psychologist*, Vol. 14 No. 1, pp. 8-17.

Hu, F.B., Goldberg, J., Hedeker, D., Flay, B.R. and Pentz, M.A. (1998), "Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes", *American Journal of Epidemiology*, Vol. 147 No. 7, pp. 694-703.

Jindal, N. and Liu, B. (2007), "Review spam detection", *Proceedings of the 16th International World Wide Web Conference*, pp. 1189-1190.

John, O.P. and Srivastava, S. (1999), "The big five trait taxonomy: history, measurement, and theoretical perspectives", in Pervin, L.A. and John, O.P. (Eds), *Handbook of Personality: Theory and Research*, Guilford Press, New York, NY, pp. 102-138.

John, O.P., Robinson, E.D. and Kentle, R.L. (1991), *The Big Five Inventory-Versions 4a and 54*, University of CA Berkeley, Institute of Personality and Social Research, Berkeley, CA.

Kannan, P.K. and Li, H.A. (2017), "Digital marketing: a framework, review and research agenda", *International Journal of Research in Marketing*, Vol. 34 No. 1, pp. 22-45.

Kanouse, D.E. and Hanson, L.R. (1972), "Negativity in evaluations", in Jones, E.E. (Ed.), *Attribution: Perceiving the Causes of Behavior*, General Learning Press, Morristown, NJ, pp. 1-26.

Kim, S., Chang, H., Lee, S., Yu, M. and Kang, J. (2015), "Deep semantic frame-based deceptive opinion spam analysis", *International Conference on Information and Knowledge Management, Proceedings*, Vol. 19, 23-Oct-2015, Association for Computing Machinery, pp. 1131-1140.

Kronrod, A., Lee, J. and Gordeliy, I. (2017), "20-F: how people compose and detect insincere product reviews: combining experimental and automated-text-analysis approaches", ACR North American Advances.

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*, Vol. 18 No. 3, pp. 185-195.

Lalonde, T.L., Nguyen, A.Q., Yin, J., Irimata, K. and Wilson, J.R. (2013), "Modeling correlated binary outcomes with time-dependent covariates", *Journal of Data Science*, Vol. 11 No. 4, pp. 715-738.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D. and Gorski, S. (2011), "Peer to peer lending: the relationship between language features, trustworthiness, and persuasion success", *Journal of Applied Communication Research*, Vol. 39 No. 1, pp. 19-37.

Levine, T.R. (2014), "Truth-default theory (TDT): a theory of human deception and deception detection", *Journal of Language and Social Psychology*, Vol. 33 No. 4, pp. 378-392.

Levine, T.R., Park, H.S. and McCornack, S.A. (1999), "Accuracy in detecting truths and lies: documenting the 'veracity effect'", *Communication Monographs*, Vol. 66 No. 2, pp. 125-144.

Li, J. and Chignell, M. (2010), "Birds of a feather: how personality influences blog writing and reading", *International Journal of Human-Computer Studies*, Vol. 68 No. 9, pp. 589-602.

Luca, M. and Zervas, G. (2016), "Fake it till you make it: reputation, competition, and yelp review fraud", *Management Science*, Vol. 62 No. 12, pp. 3412-3427.

Mafael, A., Gottschalk, S. and Kreis, H. (2016), "Examining biased assimilation of brand-related online reviews", *Journal of Interactive Marketing*, Vol. 36, pp. 91-106.

Majumdar, A. and Bose, I. (2018), "My words for your pizza: an analysis of persuasive narratives in online crowdfunding", *Information & Management*, Vol. 55 No. 6, pp. 781-794.

Mayzlin, D., Dover, Y. and Chevalier, J. (2014), "Promotional reviews: an empirical investigation of online review manipulation", *American Economic Review*, Vol. 104 No. 8, pp. 2421-2455.

Mukherjee, A., Liu, B. and Glance, N. (2012), "Spotting fake reviewer groups in consumer reviews", *Proceedings of the 21th International World Wide Web Conference*, pp. 191-200.

Munzel, A. (2015), "Malicious practice of fake reviews: experimental insight into the potential of contextual indicators in assisting consumers to detect deceptive opinion spam", *Recherche et Applications En Marketing (English Edition)*, Vol. 30 No. 4, pp. 24-50.

Ngo-Ye, T., Sinha, A. and Sen, A. (2016), "Predicting the helpfulness of online reviews using a scripts-enriched text regression model", *Expert Systems with Applications*, Vol. 71, pp. 98-110.

Ott, M., Cardie, C. and Hancock, J.T. (2012), "Estimating the prevalence of deception in online review communities", *Proceedings of International World Wide Web Conference*, pp. 201-210.

Ott, M., Cardie, C. and Hancock, J.T. (2013), "Negative deceptive opinion spam", *Proceedings of North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2013)*, pp. 497-501.

Ott, M., Choi, Y., Cardie, C. and Hancock, J.T. (2011), "Finding deceptive opinion spam by any stretch of the imagination", *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1, 309- 319.

Pan, W. (2001), "Akaike's information criterion in generalized estimating equations", *Biometrics*, Vol. 57 No. 1, pp. 120-25.

Pan, L. and Chiou, J.-S. (2011), "How much can you trust online information? Cues for perceived trustworthiness of consumer-generated online information", *Journal of Interactive Marketing*, Vol. 25 No. 2, pp. 67-74.

Park, C. and Lee, T. (2009), "Information direction, website reputation and eWOM effect: a moderating role of product type", *Journal of Business Research*, Vol. 62 No. 1, pp. 61-67.

Pavlou, P.A. and Dimoka, A. (2006), "The nature and role of feedback text comments in online marketplaces: implications for trust building, price premiums, and seller differentiation", *Information Systems Research*, Vol. 17 No. 4, pp. 392-414.

Pennebaker, J.W. and King, L.A. (1999), "Linguistic styles: language use as an individual difference", *Journal of Personality and Social Psychology*, Vol. 77 No. 6, pp. 1296-1312.

Pennebaker, J.W., Boyd, R.L., Jordan, K. and Blackburn, K. (2015), *The Development and Psychometric Properties of LIWC2015*, University of Texas at Austin, Austin, TX.

Petty, R.E. and Cacioppo, J.T. (1981), *Attitudes and Persuasion: Classic and Contemporary Approach*, William C. Brown, Dubuque, IA.

Petty, R.E. and Cacioppo, J.T. (1986), "The elaboration likelihood model of persuasion", *Advances in Experimental Social Psychology*, pp. 124-125.

Reimer, T. and Benkenstein, M. (2016), "When good WOM hurts and bad WOM gains: the effect of untrustworthy online reviews", *Journal of Business Research*, Vol. 69 No. 12, pp. 5993-6001.

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*, Vol. 21 No. 4, pp. 76-94.
Taraban, R. and Khaleel, A. (2019), "Analyzing topic differences, writing quality, and rhetorical context in college students' essays using Linguistic Inquiry and Word Count", *East European Journal of Psycholinguistics*, Vol. 6 No. 2, pp. 107-118.
Toma, C.L. and D'Angelo, J.D. (2015), "Tell-tale words: linguistic cues used to infer the expertise of online medical advice", *Journal of Language and Social Psychology*, Vol. 34 No. 1, pp. 25-45.
Toma, C.L. and Hancock, J.T. (2012), "What lies beneath: the linguistic traces of deception in online dating profiles", *Journal of Communication*, Vol. 62 No. 1, pp. 78-97.
Wang, R.Y. and Strong, D.M. (1996), "Beyond accuracy: what data quality means to data consumers", *Journal of Management Information Systems*, Vol. 12 No. 4, pp. 5-34.
Wathen, C.N. and Burkell, J.A. (2002), "Believe it or not: factors influencing credibility on the web", *Journal of the American Society for Information Science and Technology*, Vol. 53 No. 2, pp. 134-144.
Wells, G. and Alpert, L.I. (2018), "In Facebook's effort to fight fake news, human fact-checkers struggle to keep up", *The Wall Street Journal*, available at: www.wsj.com
Xiao, B. and Benbasat, I. (2011), "Product-related deception in e-commerce: a theoretical perspective", *MIS Quarterly*, Vol. 35 No. 1, pp. 169-195.
Yoo, K.H. and Gretzel, U. (2009), *Comparison of Deceptive and Truthful Travel Reviews. Information and Communication Technologies in Tourism 2009*, Springer, Vienna.
Zablocki, A., Makri, K. and Houston, M. (2019), "Emotions within online reviews and their influence on product attitudes in Austria, USA and Thailand", *Journal of Interactive Marketing*, Vol. 46, pp. 20-39.
Zhu, Y., Liu, M., Zeng, X. and Huang, P. (2020), "The effects of prior reviews on perceived review helpfulness: a configuration perspective", *Journal of Business Research*, Vol. 110, pp. 484-494.
Zuckerman, M., DePaulo, B.M. and Rosenthal, R. (1981), "Verbal and nonverbal communication of deception", in Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, Vol. 14, Academic Press, New York, NY, pp. 1-59.

Appendix 1 Sample reviews on Monaco Hotel Chicago

Negative fake

My Family and I went to this hotel on holiday last month. It was a very bad hotel. When we got there we imediatly checked in and went up to our room. The view from the window was a brick wall. Really? We even asked the Lasy at the front desk, but she said they couldnt put us in a room with a view cause we didn!t book one and we couldn't upgrade. Also, My Wife is a light sleeper and the trains kept her up at night. Another thing is alot of ads boast a hot tub in all the rooms but this is not true, our room only had a normal tub. But, the tub had no stopper, so we could only take showers. The room was also very small compared to other hotels with around the same prices.

Negative real

I am staying here now and actually am compelled to write this review before I fall asleep. The front desk staff were brief and one of them was chatting with her friend (gossiping) as I checked in (VERY unprofessional). The room she offered me was on a 'high' floor (5th [...]. HA) and when I checked in there was a big bag of grapes left behind by the previous guest on the window sill and some of the previous guests hair in the bathtub [...].l was DONE. Called down and had them re-clean the room; they made no offer to compensate or upgrade [...] not even a bottle of wine. The front desk clerk suggested I come down and get a key to another room (like I have time for that). Disappointing as I do like Kimpton hotels and am a Kimpton In Touch member (btw theyve done away with the amenities [...] you just get a free item from the mini bar).

Positive fake

My stay at the Hotel Monaco Chicago was amazing. The staff are polite and well poised, eager to give a helping hand in a short notice. To someone like me, who had never even been in the city of Chicago, it really gave me a good feeling throughout my whole trip. The rooms, hallways and facilities were exceptionally clean and tidy, and whenver I went out, I would always find my room perfect; be it for a night of undwinding after one of my conferences, or just to hit the mattress and sleep. During my stay, I stopped at their restuarant where I had one of the best American style meal in a while. Overall, the Hotel Monaco is a place I would surely stay at again if given the chance to visit Chicago for a second time. It is truely exceptional.

Positive real

After reading so many great reviews I booked the Hotel Monaco and was not disappointed. I booked through their web site about 2 1/2 months prior to arrival for S170/night plus tax for a room with 2 queens. The location was perfect – right by the river near Michigan Ave. - minutes walk to the El, Millenium Park, the Theater District and Shopping. The room was clean, a little larger than standard and nicely decorated. We had window seats in the room with a view of the river. There was Starbucks coffee in the lobby in the morning and a wine reception in the evening. To top it off, the staff was very friendly and knowledgable about the area.

Appendix 2

Table A1 Big 5 Inventory measurement means, standard deviations and alpha reliability coefficients

Big Five Inventory [John *et al.* (1991)]
(5 points, 1 = strongly disagree, 5 = strongly agree; Min = –2, Max = +2)
Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please indicate the extent to which you agree or disagree with each of the following statements. I see myself as someone who. . .

***Extraversion* (*M* = –0.09, SD = 0.84, α = 0.89)**

1	Is talkative
2	Is reserved (R)
3	Is full of energy
4	Generates a lot of enthusiasm
5	Tends to be quiet (R)
6	Has an assertive personality
7	Is sometimes shy, inhibited (R)
8	Is outgoing, sociable

***Neuroticism* (*M* = –0.14, SD = 0.88, α = 0.90)**

1	Is depressed, blue
2	Is relaxed, handles stress well (R)
3	Can be tense
4	Worries a lot
5	Is emotionally stable, not easily upset (R)
6	Can be moody
7	Remains calm in tense situations (R)
8	Gets nervous easily

***Conscientiousness* (*M* = 0.74, SD = 0.62, α = 0.85)**

1	Does a thorough job
2	Can be somewhat careless (R)
3	Is a reliable worker
4	Tends to be disorganized (R)
5	Tends to be lazy (R)
6	Perseveres until the task is finished
7	Does things efficiently
8	Makes plans and follows through with them
9	Is easily distracted (R)

***Agreeableness* (*M* = 0.69, SD = 0.62, α = 0.82)**

1	Tends to find fault with others (R)
2	Is helpful and unselfish with others
3	Starts quarrels with others (R)
4	Has a forgiving nature
5	Is generally trusting
6	Can be cold and aloof (R)
7	Is considerate and kind to almost everyone
8	Is sometimes rude to others (R)
9	Likes to cooperate with others

***Openness* (*M* = 0.65, SD = 0.61, α = 0.83)**

1	Is original, comes up with new ideas
2	Is curious about many different things
3	Is ingenious, a deep thinker
4	Has an active imagination
5	Is inventive
6	Values artistic, aesthetic experiences
7	Prefers work that is routine (R)
8	Likes to reflect, play with ideas
9	Has few artistic interests (R)
10	Is sophisticated in art, music, or literature

Note: "R" denotes reverse-scored items

Journal of Consumer Marketing
Volume 39 · Number 5 · 2022 · 523–537

Appendix 3

Table A2 Effects of review and consumer characteristics on identification accuracy

Variable	Model A Positive real *B* (*Z*)	Model B Positive fake *B* (*Z*)	Model C Negative real *B* (*Z*)	Model D Negative fake *B* (*Z*)
Intercept	–0.656 (4.91)***	–0.392 (4.16)***	–0.368 (3.61)***	–1.497 (6.38)***
Review characteristics				
Word count	–0.035 (0.77)	–0.048 (1.11)	0.003 (0.14)	0.155 (2.99)***
Words per sentence	–0.056 (1.89)**	–0.026 (0.43)	0.060 (2.49)***	0.058 (0.47)
Causation	–0.054 (1.51)	–0.009 (0.36)	0.049 (2.08)**	0.120 (1.96)**
Affect	–0.006 (0.18)	–0.037 (1.28)	–0.017 (0.61)	0.097 (1.10)
Consumer characteristics				
Big Five personality				
Neuroticism	–0.007 (0.18)	0.006 (0.20)	0.030 (1.01)	–0.119 (1.59)
Extraversion	–0.037 (0.97)	–0.038 (1.36)	–0.022 (0.77)	–0.054 (0.80)
Agreeableness	0.034 (0.86)	–0.059 (2.15)**	0.026 (0.86)	0.059 (0.79)
Openness	0.059 (1.54)	0.054 (1.99)**	0.012 (0.43)	–0.016 (0.24)
Conscientiousness	–0.023 (0.65)	–0.007 (0.25)	0.018 (0.58)	–0.014 (0.19)
Demographics				
Gender	0.082 (1.13)	0.015 (0.27)	0.026 (0.48)	0.005 (0.04)
Age	–0.046 (1.59)	–0.012 (0.57)	–0.034 (1.48)	0.089 (1.94)**
Education	–0.027 (0.80)	0.009 (0.35)	0.012 (0.44)	–0.007 (0.12)
Control variable				
Prior familiarity	0.049 (1.49)	0.049 (2.14)**	0.021 (0.79)	–0.011 (0.16)
QIC	821.120	818.998	821.475	826.045

Notes: Dependent variable: consumer's identification accuracy (not accurate= 0, accurate = 1); gender (male = 0, female = 1); $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Corresponding author

Shabnam Azimi can be contacted at: sazimi@luc.edu