

# Theorizing the textual differences between authentic and fictitious reviews: Validation across positive, negative and moderate polarities

Banerjee, Snehasish; Chua, Alton Yeow Kuan

2017

Banerjee, S., & Chua, A. Y. K. (2017). Theorizing the textual differences between authentic and fictitious reviews: Validation across positive, negative and moderate polarities. *Internet Research*, 27(2), 321-337.

<https://hdl.handle.net/10356/84164>

<https://doi.org/10.1108/IntR-11-2015-0309>

---

© 2017 Emerald Publishing Limited. This is the author created version of a work that has been peer reviewed and accepted for publication by Internet Research, Emerald Publishing Limited. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<http://dx.doi.org/10.1108/IntR-11-2015-0309>].

*Downloaded on 23 Apr 2021 15:36:31 SGT*

# **Theorizing the Textual Differences between Authentic and Fictitious Reviews: Validation across Positive, Negative and Moderate Polarities**

Snehasish Banerjee and Alton Y.K. Chua

Wee Kim Wee School of Communication and Information

Nanyang Technological University, Singapore

## **Abstract**

**Purpose:** The purpose of this paper is two-fold: (1) to build a theoretical model that identifies textual cues to distinguish between authentic and fictitious reviews, and (2) to empirically validate the theoretical model by examining reviews of positive, negative as well as moderate polarities.

**Design/methodology/approach:** Synthesizing major theories on deceptive communication, the theoretical model identifies four constructs—comprehensibility, specificity, exaggeration and negligence—to predict review authenticity. The predictor constructs were operationalized as holistically as possible. To validate the theoretical model, 1,800 reviews (900 authentic + 900 fictitious) evenly spread across positive, negative and moderate polarities were analyzed using logistic regression.

**Findings:** The performance of the proposed theoretical model was generally promising. However, it could better discern authenticity for positive and negative reviews compared with moderate entries.

**Originality/Value:** The paper advances the extant literature by theorizing the textual differences between authentic and fictitious reviews. It also represents one of the earliest attempts to examine nuances in the textual differences between authentic and fictitious reviews across positive, negative as well as moderate polarities.

**Keywords:** user-generated content, online review, fake review, text analysis, linguistic analysis, authenticity, deception, sentiment

**Article Classification:** Research paper

## 1. Introduction

With the proliferation of web stores and online shopping, consumers increasingly read user-generated online reviews prior to purchase (Al-Debei *et al.*, 2015; Owusu *et al.*, 2016). Unlike marketer-generated information that tends to be worded positively, the polarity of reviews varies substantially (Jeong and Koo, 2015; Kugler, 2014). Concurrently, some consumers seek extreme positive and negative reviews to confirm or eliminate purchase options (Mackiewicz, 2015). Yet, others seek moderate reviews that appear more realistic than extreme ones (Lee and Yang, 2015).

Regardless of polarity however, not all reviews describe authentic post-purchase experiences. As the popularity of reviews charts a meteoric rise, submitting fictitious entries is becoming a common malpractice. Businesses heap praises and criticisms in fictitious reviews to gain an unfair edge over their rivals (Kugler, 2014). Users could post positive and negative fictitious reviews simply for fun, or out of boredom. To enhance credibility, fictitious reviews could also maintain a moderate tone to sound realistic. Given the variety of reviews available, consumers find it challenging to discern what is authentic, and what is fictitious.

In response to the challenge, scholars mostly use text analysis to distinguish between authentic and fictitious reviews (Harris, 2012; Yoo and Gretzel, 2009). However, two research gaps are identified. First, the literature lacks a comprehensive model underpinned by relevant theories to study textual cues as predictors of review authenticity. In the absence of

adequate theorizing, scholars appear unsure about the width of the net to cast in order to identify textual differences between authentic and fictitious reviews. For example, while some narrowly analyzed textual cues such as comprehensibility and writing style (Yoo and Gretzel, 2009), others expanded the set to include the level of specific details (Banerjee and Chua, 2014). At the extreme end, Ott *et al.* (2011) used almost a brute-force strategy to include a large array of textual cues ranging from psycholinguistic dimensions to n-grams, thereby compromising parsimony. Hence, theorizing the textual differences between authentic and fictitious reviews is needed.

Second, little empirical research has hitherto examined how authentic reviews textually differ from fictitious ones across positive, negative as well as moderate polarities. Reviews with a positive polarity recommend products and services, while those with a negative slant offer criticisms. Unlike these acute polarities, a moderate polarity is used to express mild opinions. The acuteness of the positive and the negative polarities may make authentic reviews substantially different from fictitious entries (Ott *et al.*, 2013; Yoo and Gretzel, 2009). However, the mildness of the moderate polarity may blur the lines between the two (Wu *et al.*, 2010). This assumption highlights the need for finer-grained analysis across polarities.

For these reasons, the purpose of this paper is two-fold. First, it builds a theoretical model that identifies textual cues to distinguish between authentic and fictitious reviews. Second, the paper attempts to empirically validate the theoretical model by examining reviews of positive, negative as well as moderate polarities.

The paper makes two-fold contributions. First, it advances the theoretical understanding of textual differences between authentic and fictitious reviews. Writing fictitious reviews is a form of deceptive communication that involves transmitting fictitious messages by resembling authentic ones to foster a false conclusion by receivers. Therefore,

this paper synthesizes four major deceptive communication theories—the information manipulation theory, the leakage theory, the self-presentational theory, and the reality monitoring theory—in the context of reviews.

Second, motivated by the assumption that textual differences between authentic and fictitious reviews could be disparate across polarities, this paper represents one of the earliest attempts to examine review authenticity across positive, negative as well as moderate entries. By highlighting subtle nuances across the polarities, this paper explores an uncharted territory. This facilitates gleaning new insights for practice.

The paper proceeds as follows. Section 2 reviews the related works, which is followed by the proposed theoretical model in Section 3. The methods are described in Section 4. Section 5 presents the results, which are further discussed in Section 6.

## **2. Related Works**

Studies that seek to distinguish between authentic and fictitious reviews are aplenty. Some studies relied on metadata. For example, Li *et al.* (2011) used review-related and reviewer-related metadata as predictors of authenticity while Jindal and Liu (2008) additionally considered product-related metadata.

Some studies focused on patterns to distinguish between authentic and fictitious reviews. For example, studies such as Wu *et al.* (2010) suggested that positive reviews that are posted soon after a commodity has attracted a series of negative entries are likely to be fictitious. Again, studies such as Carbunar *et al.* (2013) conceived fictitious reviews as those that distort the natural distribution of review ratings attracted by a commodity.

Another viable approach to distinguish between authentic and fictitious reviews involves relying on textual differences (Heydari *et al.*, 2015). In one of the earliest works, Yoo and Gretzel (2009) analyzed only positive reviews using textual cues such as complexity

and use of emotions. Authentic reviews were found to be less complex and emotional than fictitious entries.

The initial study triggered more scholarly efforts in the later years. Consistent with Yoo and Gretzel (2009), Wu *et al.* (2010) indicated that positive and negative reviews that expressed extreme emotions were more likely to be fictitious compared with those that indicated moderate opinions. Ong *et al.* (2014) further echoed Yoo and Gretzel's (2009) findings by confirming that authentic reviews were less complex than fictitious ones when considering positive polarity only. Counter-intuitively, Banerjee and Chua (2014) found that authentic reviews were less complex yet more difficult to be read compared with fictitious ones. Moreover, when a similar approach was used to study both positive and negative reviews, Harris (2012) found authentic reviews to be more complex than fictitious entries.

A separate group of studies emphasized on accurate classification of authentic and fictitious reviews without necessarily identifying the ways in which the two were different. In such a study that analyzed reviews of only positive polarity, Ott *et al.* (2011) expanded the scope of investigation to a large array of textual cues. These comprised level of specific details, psycholinguistic dimensions that included cues of exaggeration and negligence, along with n-grams. A high classification accuracy was obtained. Using a similar approach, Ott *et al.* (2013) found high classification accuracy for negative reviews too.

Despite these efforts, little scholarly investigation has hitherto been trained on review authenticity for moderate reviews. This calls for not only building a theoretical model of textual cues to predict review authenticity but also validating such a model using reviews of positive, negative as well as moderate polarities.

### **3. Theoretical Model**

This paper develops a theoretical model by synthesizing four major deceptive communication theories—the information manipulation theory, the leakage theory, the self-presentational theory, and the reality monitoring theory. These were selected because they collectively help identify textual cues to distinguish between authentic and fictitious reviews.

Specifically, the information manipulation theory posits that authentic messages differ from fictitious ones based on the maxims of quantity, quality, relation and manner (McCornack, 1992). Quantity is the amount of information in a message, quality is its extent of details, relation refers to its relevance, and manner indicates its writing style. The leakage theory posits that authentic messages differ from fictitious ones based on the extent to which the entries coincidentally leak out clues—often out of negligence—that help deception detection (Ekman and Friesen, 1969). The self-presentational theory posits that authentic messages differ from fictitious ones based on the extent to which they are written with the motivation to appear favorable (DePaulo *et al.*, 2003). The reality monitoring theory posits that authentic messages differ from fictitious ones based on the amount of perceptual details, contextual details, and indicators of cognition (Johnson and Raye, 1981). Perceptual details are commonly expressed using sensory words, contextual details indicate references to space or time, and indicators of cognition include words that indicate psychical processing.

Other related theories such as the interpersonal deception theory (Buller and Burgoon, 1996) and the management obfuscation theory (Bloomfield, 2002) were excluded because they were deemed inappropriate for online reviews. The interpersonal deception theory is used to study deception in dialogic communication while the management obfuscation theory is specifically suited to study deception in businesses' annual reports, which mostly comprise numbers and charts.

Guided by Weber (2012), this paper conceives the task of theory-building as one that logically combines constructs, associations and boundaries. Constructs refer to concepts,

associations are the relationships among constructs, and boundaries refer to a theory's delimitation (McLeod and Pan, 2005). Constructs and associations in the proposed theoretical model are synthesized from the four major deception communication theories. In particular, the theoretical model identifies four predictor constructs—comprehensibility, specificity, exaggeration and negligence of reviews—that are hypothesized to be associated with the outcome construct authenticity as shown in Figure 1. With respect to boundary, the scope of the theoretical model is delimited by the definition of its outcome construct—authenticity. It treats authenticity as a proxy for real experiences. In other words, even though it is applicable to distinguish between experiential and imaginative texts, it is beyond its scope to distinguish between truths and lies.

Comprehensibility of a review refers to the extent to which it is easy to understand. The association between comprehensibility and authenticity stems from the information manipulation theory, and the self-presentational theory. The information manipulation theory suggests that authentic and fictitious messages are distinguishable based on their quantity and manner (McCornack, 1992), which in turn shape the extent to which the entries are easy to comprehend. Moreover, the self-presentational theory suggests that individuals could self-present by using sophisticated language to write reviews convincingly. Specifically, those writing fictitious reviews could go overboard in using sophisticated language, thereby making the entries difficult to comprehend (DePaulo *et al.*, 2003). Hence, the following is hypothesized:

H1: Comprehensibility of reviews helps to ascertain their authenticity.

Specificity of a review refers to its richness in details. The association between specificity and authenticity stems from the information manipulation theory, and the reality monitoring theory. The information manipulation theory suggests that authentic and fictitious

messages are distinguishable based on their quality and relevance (McCornack, 1992), which in turn shape the extent to which the entries are rich in details. Moreover, the reality monitoring theory suggests that authentic and fictitious messages could be distinguished based on their use of perceptual and contextual details (Johnson and Raye, 1981), which are also known to determine the extent to which the entries are detailed. Hence, the following is hypothesized:

H2: Specificity of reviews helps to ascertain their authenticity.

Exaggeration of a review refers to the writing style used in the entry to sound convincing. The association between exaggeration and authenticity stems from the leakage theory, and the self-presentational theory. Based on the leakage theory (Ekman and Friesen, 1969), it seems that even though individuals writing authentic messages and those writing fictitious ones could both try to sound convincing, the latter might inadvertently exaggerate to prove a point. Moreover, the self-presentational theory suggests that individuals writing fictitious reviews could inadvertently rely on the rhetoric of exaggeration in attempt to prove their credence and sound convincing (DePaulo *et al.*, 2003). Hence, the following is hypothesized:

H3: Exaggeration of reviews helps to ascertain their authenticity.

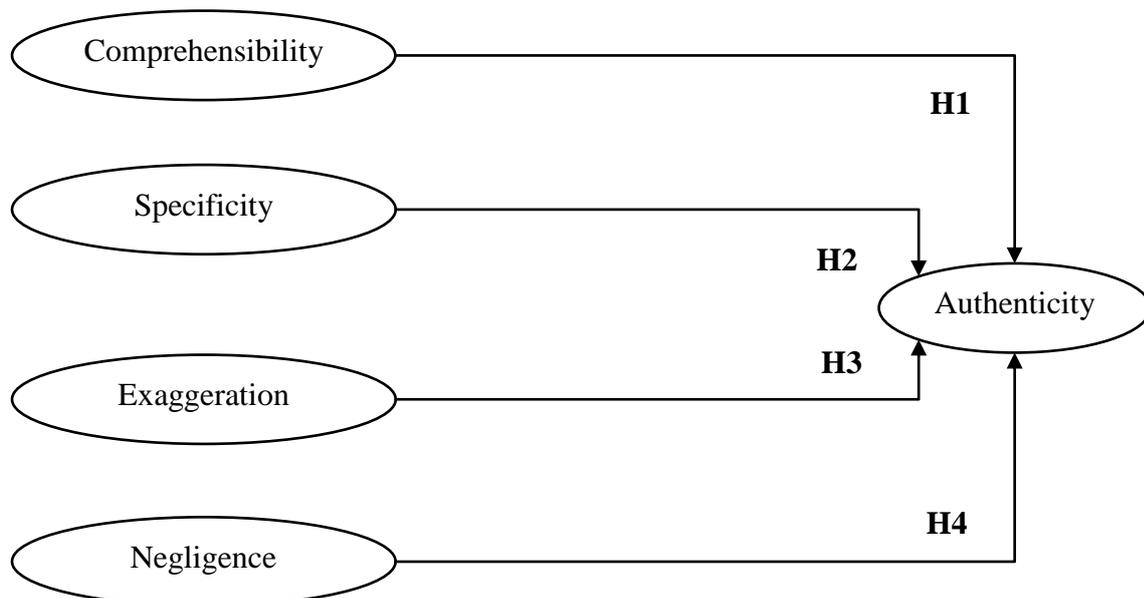
Negligence of a reviews refers to clues for deception detection that are coincidentally leaked out in the entry. The association between comprehensibility and authenticity stems from the leakage theory, and the reality monitoring theory. The leakage theory suggests that since writing fictitious messages based on imagination is cognitively challenging, individuals could be negligent in performing this task. In consequence, authentic messages could differ from fictitious ones based on the extent to which the entries coincidentally leak out clues that

help deception detection (Ekman and Friesen, 1969). Moreover, the reality monitoring theory suggests that authentic and fictitious messages could be distinguished based on the extent of psychological processing engaged by their respective contributors. The extent of psychological processing has the potential to shape the level of negligence in messages (Johnson and Raye, 1981). Hence, the following is hypothesized:

H4: Negligence of reviews helps to ascertain their authenticity.

The associations between the predictor constructs and the outcome construct in the four hypotheses are phrased non-directionally in view of the lack of scholarly consensus among related works. For example, some studies found that authentic reviews were easier to comprehend compared with fictitious ones (Yoo and Gretzel, 2009) whereas others found the converse to be true (Harris, 2012). While some indicated authentic reviews to be more exaggerated than fictitious entries (Yoo and Gretzel, 2009), others suggested the possibility of blurring such differences (Banerjee and Chua, 2014). The debates in this area of research advocated caution in presenting directional hypotheses.

**[Insert Figure 1 here]**



**Figure 1.** The proposed theoretical model

## **4. Methods**

### *4.1. Data Collection*

This paper uses hotel reviews for investigation because reviews are crucial for high-credence services such as those related to tourism (Jalilvand and Samiei, 2012; Pan *et al.*, 2013). Authentic reviews refer to those contributed by users who had paid and patronized a given hotel whereas fictitious reviews refer to those written by individuals based on imagination without any experience of staying in the hotel.

To cater to these definitions, authentic reviews were collected from authenticated review websites that solicit entries only from bona fide travellers after their purchases and stays. Several recent studies suggest that it is unlikely for spammers to purchase a commodity only in order to be able to post a fictitious review (Crawford *et al.*, 2015; Ong *et al.*, 2014). On the other hand, fictitious reviews were solicited from participants in a research setting. Participants were requested not to write reviews for hotels where they had earlier stayed.

Three popular authenticated review websites largely similar in functionalities were chosen, namely, Agoda.com, Expedia.com and Hotels.com. All three routinely collect reviews from bona fide travellers by sending them unique URLs after they had checked out. Reviews are required to contain two parts: titles and descriptions. They are made available online after being vetted by website administrators.

A list of 15 hotels in Asia (Table I) that had attracted almost 1,000 or more reviews across the three chosen websites were randomly identified on 28 February, 2013. The identified hotels uniformly straddled across five popular tourist destinations in Asia, namely, Bangkok, Hong Kong, Kuala Lumpur, Singapore, and Tokyo, as well as three hotel types—

luxury, budget and mid-range. Hotel types were ascertained by checking the consistency of their star categories across the chosen websites.

**[Insert Table I here]**

For each of the 15 hotels, 60 authentic reviews (20 positive + 20 negative + 20 moderate) were randomly collected to yield 900 authentic reviews altogether (300 positive + 300 negative + 300 moderate). All reviews were in English, contained meaningful titles, and meaningful descriptions of minimally 150 characters.

Reviews were categorized as positive, negative or moderate based on the polarity of their ratings (Chen and Lurie, 2013). Specifically, Expedia.com and Hotels.com use a 5-point rating scale. For reviews from these platforms, reviews with ratings of one or two stars were treated as negative, those with three stars were taken as moderate, and entries with four or five stars were deemed as positive (Chen and Lurie, 2013). However, Agoda.com employs a 10-point rating scale. Scales differing in range could not be linearly interpolated because those with more options yield higher scores. To make ratings from the three websites comparable, rescaling was done based on Dawes (2002).

For each of the 15 hotels, 60 fictitious reviews were solicited from more than some 400 participants, who were recruited via convenience sampling and snowballing on meeting two demographic-related eligibility criteria. First, their age had to range from 21 to 45 years. Second, in terms of their educational profile, they must minimally be undergraduate students. In other words, the participants could include undergraduate students, graduate students, or working adults with minimally an undergraduate degree. These demographic-related criteria were chosen because reviews are mostly written by young and educated individuals (Munar and Jacobsen, 2014). Additionally, the participants were required to be familiar with the use of reviews, and have travel experience during the previous year.

Participants were given the website of the hotel for which fictitious review was solicited. However, they were requested not to submit an entry had they stayed in the hotel before. There was only one instance in which a participant explicitly notified one of the authors of having stayed in a hotel for which a fictitious review was solicited. Following the request, the participant did not submit an entry for that property.

The detailed instructions given to participants to write fictitious reviews—either positive, or negative, or moderate—were informed by prior studies (Ott *et al.*, 2011; Yoo and Gretzel, 2009). Participants could choose to write at most six fictitious reviews for six different hotels. Finally, a total of 900 fictitious reviews (300 positive + 300 negative + 300 moderate) written by 284 participants (aged 21-25 years: 88, aged 26-35 years: 146, aged 36-45 years: 50; educational background: minimally undergraduate students; gender: 128 females, 156 males) were admitted for analysis. All the reviews were in English, contained meaningful titles, and meaningful descriptions of minimally 150 characters.

To ensure the validity of categorizing the 1,800 reviews (900 authentic + 900 fictitious) as positive, negative or moderate, three annotators were recruited. They had adequate experience of reading reviews. Each annotator was randomly assigned one-third of the total set of reviews. The annotators were shown the titles and the descriptions of reviews. Hotel names were concealed to avoid biases. They were asked to predict the review type as either positive or negative or moderate. Cohen's kappa for the agreement between review polarities predicted by annotators, and those based on the original categorization of ratings was at a non-chance level:  $\kappa = 0.91$  (Cohen, 1960).

#### *4.2. Construct Operationalization*

To empirically validate the theoretical model built in this paper, it was necessary to operationalize its constructs. Authenticity (the outcome construct) had been operationalized

by the definitions of authentic and fictitious reviews as indicated in Section 4.1. The four predictor constructs are operationalized as follows.

Comprehensibility includes the following: readability, word familiarity, and surface-level characteristics (Cao *et al.*, 2011; Chall and Dale, 1995; Ghose and Ipeirotis, 2011). Readability is commonly measured using indicators such as Flesch-Kincaid Grade Level, Gunning-Fog Index, Automated-Readability Index, Coleman-Liau Index, Lasbarhets Index, and Rate Index (Ghose and Ipeirotis, 2011). To recognize the strengths of each, readability is often calculated as the mean of these indicators (Chua and Banerjee, 2015; Ghose and Ipeirotis, 2011). Lower means indicate more readable reviews and vice-versa. Word familiarity could be measured by comparing every word in a text against Dale-Chall's lexicon of familiar words (Chall and Dale, 1995). Commonly studied surface-level characteristics include number of characters per word, number of words, fraction of words containing 10 or more characters (henceforth, long words), and number of words per sentence (Cao *et al.*, 2011).

Specificity includes the following: informativeness, perceptual details, contextual details, lexical diversity, and the use of function words (Ott *et al.*, 2011; Rayson *et al.*, 2001; Yoo and Gretzel, 2009). Informativeness is commonly measured based on the use of eight part-of-speech (POS) tags: nouns, adjectives, prepositions, articles, conjunctions, verbs, adverbs, and pronouns. Informative texts are richer in the first four but limited in terms of the remainder (Rayson *et al.*, 2001). Among pronouns, both singular and plural self-references need to be taken into account to glean richer insights (Boals and Klein, 2005; Hancock *et al.*, 2007). Perceptual details include proportion of visual, aural and feeling words, while contextual details comprise that of temporal and spatial words in reviews. Lexical diversity is defined as the fraction of unique words in reviews, while function words comprise vague

non-content words that add little meaning (Hancock *et al.*, 2007; Tausczik and Pennebaker, 2010).

Exaggeration includes the following: the use of affect, tense and rhetorical devices (Banerjee and Chua, 2014; Newman *et al.*, 2003; Tausczik and Pennebaker, 2010; Yoo and Gretzel, 2009; Zhang and Peng, 2015). Affect could be measured in terms of the fraction of emotion words (henceforth, emotiveness)—both positive and negative. Tense indicates the chronological focus of reviews through differential use of words in past, present and future tenses (Tausczik and Pennebaker, 2010). Informed by prior studies, the rhetorical devices considered in this paper include the proportion of firm words such as “always” and “never”, upper case characters, references to the hotel names (henceforth, brand references), as well as punctuations such as ellipses, exclamation points, and question marks (Tausczik and Pennebaker, 2010; Yoo and Gretzel, 2009).

Negligence includes the following: cues leaked due to lack of conscientiousness, and those leaked due to subconscious attempts to cover up. The former is manifested through increased use of discrepancy words such as “could”, fillers such as “you know”, tentative words such as “perhaps”, and motion words such as “arrive” but limited use of exclusion words such as “without” (Boals and Klein, 2005; Newman *et al.*, 2003; McQuaid *et al.*, 2015; Pasupathi, 2007; Tausczik and Pennebaker, 2010). The latter is manifested through increased use of causal words such as “because”, and insight words such as “think” (Boals and Klein, 2005; Newman *et al.*, 2003; McQuaid *et al.*, 2015; Pasupathi, 2007; Tausczik and Pennebaker, 2010).

Cumulatively, this paper identifies 43 operationalized indicators straddled across the four constructs—comprehensibility (6 indicators), specificity (17 indicators), exaggeration (13 indicators), and negligence (7 indicators). These are summarized in Table II.

**[Insert Table II here]**

### 4.3. Measurements and Analyses

The 43 operationalized indicators were measured using Stanford Parser's POS tagger (Klein and Manning, 2003), Linguistic Inquiry and Word Count (LIWC2007) toolkit (Pennebaker *et al.*, 2007), and some customized Java programs. Each measure was computed separately for titles and descriptions of all reviews.

However, mean readability (comprehensibility: measure #1), number of words per sentence (comprehensibility: measure #6), and ellipses (exaggeration: measure #10) could not be calculated for titles. For one, measures such as mean readability and number of words per sentence rely on the number of sentences in a given text. These are not meaningful for review titles that rarely contain complete sentences. Besides, ellipses in review titles yielded few occurrences in the dataset. Therefore, the predictors were measured using all the 43 measures for review descriptions, and the 40 measures (43 - 3) for review titles.

Data were analyzed using binomial logistic regression, which is used to predict a categorical outcome variable. In this paper, authenticity of reviews was the outcome variable (1 = *authentic*, 0 = *fictitious*) while the 83 textual measures (40 for titles + 43 for descriptions) were the predictors. The inter-correlations among the predictors were below the recommended threshold of 0.80, confirming no multicollinearity (Licht, 1995). The analysis was repeated thrice for positive, negative and moderate polarities separately.

## 5. Results

The logistic regression model demonstrated promising results in identifying authentic and fictitious reviews for all the three polarities. The performance was the best among negative reviews, followed by positive and moderate entries respectively (Table III). This finding bears a striking similarity with offline deception studies such as Anderson *et al.*

(1999), which found that participants mentioned more cues for deception detection during negative emotional events vis-à-vis positive ones. For moderate reviews, it was perhaps the easiest to blur the lines between authentic and fictitious entries as suggested by the lowest  $\chi^2$ , pseudo- $R^2$  and accuracy among the three polarities. The Hosmer-Lemeshow goodness-of-fit test emerged non-significant for all the three polarities, indicating adequate fitness of the model with the data for positive, negative as well as moderate reviews.

**[Insert Table III here]**

Table IV reports the odds ratios for the textual measures across titles and descriptions of positive, negative as well as moderate reviews. For brevity, measures that were non-significant across all three polarities are omitted.

Among positive reviews, titles of authentic entries were longer than fictitious ones ( $\text{Exp}(\beta)=1.46$ ). The former contained more exclusion words ( $\text{Exp}(\beta)=1.15$ ) but fewer firm words ( $\text{Exp}(\beta)=0.92$ ), exclamation points ( $\text{Exp}(\beta)=0.84$ ), and discrepancy words ( $\text{Exp}(\beta)=0.85$ ). With respect to descriptions, authentic reviews were richer in temporal words ( $\text{Exp}(\beta)=1.18$ ), negative emotion words ( $\text{Exp}(\beta)=1.60$ ), future tense ( $\text{Exp}(\beta)=1.64$ ), and exclusion words ( $\text{Exp}(\beta)=1.26$ ) than fictitious entries. The former however contained fewer singular ( $\text{Exp}(\beta)=0.78$ ) and plural self-references ( $\text{Exp}(\beta)=0.87$ ), feeling words ( $\text{Exp}(\beta)=0.66$ ), exclamation points ( $\text{Exp}(\beta)=0.85$ ), as well as discrepancy words ( $\text{Exp}(\beta)=0.79$ ).

Among negative reviews, titles of authentic entries were longer than fictitious ones ( $\text{Exp}(\beta)=1.31$ ). The former contained more verbs ( $\text{Exp}(\beta)=1.20$ ), spatial words ( $\text{Exp}(\beta)=1.04$ ), and positive emotion words ( $\text{Exp}(\beta)=1.07$ ) but fewer aural words ( $\text{Exp}(\beta)=0.88$ ), and exclamation points ( $\text{Exp}(\beta)=0.86$ ). With respect to descriptions, authentic reviews were lengthier ( $\text{Exp}(\beta)=1.02$ ) and richer in nouns ( $\text{Exp}(\beta)=1.16$ ), positive emotion words ( $\text{Exp}(\beta)=1.27$ ), future tense ( $\text{Exp}(\beta)=1.52$ ), ellipses ( $\text{Exp}(\beta)=1.47$ ), as well as causal

words ( $\text{Exp}(\beta)=1.35$ ) than fictitious entries. The former however contained fewer singular self-references ( $\text{Exp}(\beta)=0.75$ ), negative emotion words ( $\text{Exp}(\beta)=0.80$ ), firm words ( $\text{Exp}(\beta)=0.73$ ), brand references ( $\text{Exp}(\beta)=0.35$ ), and insight words ( $\text{Exp}(\beta)=0.71$ ).

Among moderate reviews, titles of authentic entries contained more nouns ( $\text{Exp}(\beta)=1.04$ ), adjectives ( $\text{Exp}(\beta)=1.02$ ), spatial words ( $\text{Exp}(\beta)=1.04$ ), temporal words ( $\text{Exp}(\beta)=1.09$ ), and positive emotion words ( $\text{Exp}(\beta)=1.05$ ) than fictitious ones. The former however contained fewer exclamation points ( $\text{Exp}(\beta)=0.85$ ), and discrepancy words ( $\text{Exp}(\beta)=0.89$ ). With respect to descriptions, authentic reviews were richer in nouns ( $\text{Exp}(\beta)=1.18$ ), and plural self-references ( $\text{Exp}(\beta)=1.19$ ) than fictitious entries. The former however contained fewer adjectives ( $\text{Exp}(\beta)=0.70$ ), adverbs ( $\text{Exp}(\beta)=0.74$ ), visual words ( $\text{Exp}(\beta)=0.72$ ), and feeling words ( $\text{Exp}(\beta)=0.70$ ).

**[Insert Table IV here]**

Overall, the differences between authentic and fictitious reviews turned out to be quite disparate across positive, negative and moderate reviews (Table V). The differences were also quite incoherent across titles and descriptions of reviews, suggesting that consumers might find it daunting to accurately predict review authenticity.

**[Insert Table V here]**

## **6. Discussion and Conclusion**

### *6.1. Discussion of the Findings*

This paper has built a theoretical model that identifies four constructs—comprehensibility, specificity, exaggeration and negligence—to predict review authenticity. The constructs were operationalized and measured as holistically as possible. Thereafter, these measures were used to examine variations in differences between authentic and fictitious reviews across positive, negative as well as moderate polarities. Discerning

authenticity turned out to be more difficult for moderate reviews vis-à-vis positive and negative entries. Some indicators corresponding to each of the predictor constructs in the theoretical model helped ascertain authenticity for positive, negative or moderate reviews (cf. Table IV). Hence, all the hypotheses (H1 through H4) were generally supported. This lends support to the robustness of the proposed theoretical model as well as its underpinning deceptive communication theories, namely, the information manipulation theory, the leakage theory, the self-presentational theory, and the reality monitoring theory.

The findings corresponding to the hypotheses (cf. Table V) are discussed more granularly in relation to the literature as follows. With respect to comprehensibility (H1), authentic and fictitious reviews could be distinguished in terms of their length, especially that of titles for positive and negative reviews. Prior research on comprehensibility of authentic and fictitious reviews has often been inconclusive (Harris, 2012; Yoo and Gretzel, 2009). This paper suggests that differences between the two are at times easily blurred. Nonetheless, titles of positive and negative authentic reviews could be lengthier than those of fictitious entries. Besides, descriptions of negative authentic reviews could be longer than their fictitious counterpart.

With respect to specificity (H2), although authentic reviews were expected to be rich in textual measures such as nouns, adjectives, spatial and temporal words with limited use of word categories such as adverbs, pronouns and function words (Ott *et al.*, 2011; Rayson *et al.*, 2001; Tausczik and Pennebaker, 2010; Yoo and Gretzel, 2009), such differences were not consistently found across the three polarities. For example, fictitious reviews were expected to contain fewer singular self-references like “I” but more plural self-references like “we” compared with authentic ones. After all, that would have allowed individuals writing fictitious reviews to dissociate themselves from the imaginary content to transfer accountability to some imaginary others (Hancock *et al.*, 2007). Contrary to expectations

however, positive and negative fictitious reviews emerged rich in singular self-references while moderate fictitious reviews contained limited plural self-references.

With respect to exaggeration (H3), authenticity could be ascertained adequately especially for positive and negative reviews. Interestingly, positive authentic reviews contained more negative emotion words than their fictitious counterpart. Conversely, negative authentic reviews contained more positive emotion words vis-à-vis negative fictitious ones. Unlike authentic reviews that seem more realistic, fictitious entries perhaps overdo the use of emotions to prove a point. The use of exclamation points in titles emerged as a significant predictor of review authenticity. For positive, negative as well as moderate reviews, titles of authentic entries contained fewer exclamation points than those of fictitious ones. Since reviews potentially influence the present image as well as the future revenues of businesses, fictitious reviews could contain fewer past tense, but more present and future tense to influence the present and the future reputation of hotels (Gunsch *et al.*, 2000; Tausczik and Pennebaker, 2010). However, positive and negative authentic reviews emerged as being rich in future tense.

With respect to negligence (H4), the differences between authentic and fictitious reviews were largely inconsistent across polarities. Positive authentic reviews contained fewer discrepancy words but more exclusion words compared with their fictitious counterpart. Negative authentic reviews contained more causal words but fewer insight words than their fictitious counterpart. The differences were largely blurred for moderate reviews. Prior research suggests that authentic reviews could contain more exclusion words but fewer discrepancy, causal and insight words (Boals and Klein, 2005; Newman *et al.*, 2003; Tausczik and Pennebaker, 2010). This paper augments such studies by demonstrating that the use of such negligence cues differs across positive, negative and moderate polarities.

## 6.2. Implications for Research

This paper offers two-fold implications for research. First, it advances the theoretical understanding of how authentic reviews textually differ from fictitious ones by proposing a theoretical model. Theory-building is considered to be the ‘jewel in the crown’ of scholarly efforts (Eisenhardt, 1989; Weber, 2012). Concurrently, the problems posed by the prevalence of both authentic and fictitious reviews are widely recognized by scholars (Ott *et al.*, 2011; Harris, 2012). Yet, the literature in this vein lacked an overarching theoretical model underpinned by relevant theories hitherto. As a result, this paper makes a timely and significant contribution by buttressing its investigation with a priori theoretical predictions.

Moreover, this paper represents the first step toward distilling the essence of major deceptive communication theories in the context of online reviews. Synthesizing four major theories, namely, the information manipulation theory (McCornack, 1992), the leakage theory (Ekman and Friesen, 1969), the self-presentational theory (DePaulo *et al.*, 2003), and the reality monitoring theory (Johnson and Raye, 1981), the paper proposed a theoretical model that identifies four constructs—comprehensibility, specificity, exaggeration and negligence—to predict the outcome construct review authenticity. Since the theoretical model includes only five constructs (four predictors and one outcome), it can be deemed parsimonious. Its empirical validation yielded generally promising results (cf. Table III), thereby confirming its robustness (Weber, 2012).

Second, this paper offers a way to holistically operationalize the textual cues of comprehensibility, specificity, exaggeration and negligence (cf. Table II). A holistic operationalization of constructs coupled with identification of their underlying dimensions is recommended to test theories precisely (McLeod and Pan, 2005). Informed by the recommendation, this paper identifies some indicators (e.g., word familiarity, lexical

diversity, use of ellipses) of the constructs that have at best received cursory attention among previous studies.

Moreover, the operationalization of some of the constructs has thus far been largely inconsistent. For example, some studies operationalized comprehensibility as average word length (Yoo and Gretzel, 2009) while others operationalized it using a readability indicator (Harris, 2012). Inadequate theorizing coupled with such fragmented operationalization of constructs in this area of research makes the literature fragmented. Therefore, this paper not only represents an attempt to synthesize the literature but also offers holistic operational definitions of constructs. These definitions could help future studies on text analysis and text mining to glean richer findings.

Third, this paper offers a methodological template to create ground truth of authentic and fictitious reviews. Creating ground truth has been an elusive problem for scholars to tackle. Some deemed duplicate reviews as fictitious and non-duplicates as authentic (Jindal and Liu, 2008). Some studies sought fictitious reviews from participants whereas authentic ones were ironically collected from unauthenticated websites (Ott et al., 2011). In this vein, this paper recommends collecting authentic reviews from authenticated websites, and soliciting fictitious reviews from participants. Alternatively, scholars could also consider drawing fictitious reviews from unauthenticated websites. This is because reviews from unauthenticated websites were found to substantially differ from entries posted in authenticated websites based on the proposed theoretical model (cf. Appendix). In other words, reviews from unauthenticated websites could be assumed to be fictitious.

### *6.3. Implications for Practice*

This paper recommends consumers to be weary of reviews that contain exclamations in titles across positive, negative as well as moderate entries. This is because it was found that

reviews with exclamations in titles were likely to be fictitious regardless of polarity. Also, consumers are advised to take reviews that describe personal experiences with a grain of salt especially for positive and negative polarities. This is because positive and negative reviews rich in singular self-references were found likely to be fictitious. Nonetheless, consumers could rely on positive and negative reviews that maintain a mild tone. This is because on the one hand, positive reviews that contained negative comments were found likely to be authentic. On the other, negative reviews that contained positive comments were also found likely to be authentic.

For review websites that deploy automated techniques to weed out fictitious from authentic reviews, this paper demonstrates that a single one-size-fits-all algorithm might not suffice. This is because this paper found discerning authenticity of positive and negative reviews to be easier compared with moderate entries. Conversely, as indicated earlier in Section 5, blurring the lines between authentic and fictitious reviews was perhaps the easiest for moderate polarity. Therefore, the results of this paper could be leveraged by review websites to calibrate their algorithms to distinguish between authentic and fictitious reviews of different polarities.

This paper also found that authentic and fictitious reviews across the three polarities were comparable in terms of numerous textual measures. This hints at two possibilities. First, fictitious reviews were written by deliberately mimicking authentic ones. Second, authentic reviews were inadvertently written in such a way that they carried traits of fictitious entries. While the first possibility is expected, the second is ominous. It suggests that individuals submitting authentic reviews do not necessarily write the entries in expected ways. This is why authentic reviews were not always informative and rich in perceptual as well as contextual details. Although authentic reviews are expected to be rich in past tense since they describe prior experiences of using a commodity, such a finding was also inconspicuous.

To counter this problem, administrators of review websites need to offer guidelines on ways to write authentic reviews. These guidelines should be disseminated to bona fide users only through personalized emails to prevent alerting potential spammers. Furthermore, measures should be taken to permit contribution of reviews only after a verified purchase. This could help minimize the prevalence of fictitious entries as much as possible, thereby offering consumers a relatively safe online shopping environment (Hsu *et al.*, 2014).

#### *6.4. Limitations and Future Research*

The findings gleaned from this paper should be viewed in light of three constraints. First, this paper analyzed differences between authentic and fictitious reviews across positive, negative and moderate comments for hotels. Also, only reviews written in English were taken into consideration. The use of such a dataset confines the generalizability of the findings. Future studies could validate the proposed theoretical model with reviews for not only other types of products and services but also encompassing those in non-English languages. Such studies are needed to theoretically extend the proposed model.

Second, the results of this paper were obtained in view of the current approaches used to write fictitious reviews. The detected textual differences between authentic and fictitious reviews could well diminish in the future, thereby falsifying the proposed theoretical model. After all, spammers might become more adept to write fictitious reviews resembling authentic entries. As the problem at hand appears to have no easy solution, this paper serves as a call for scholars to devise new approaches extending beyond textual analysis to distinguish between authentic and fictitious reviews.

Third, this paper employed a relatively covert approach in collecting authentic reviews than the one used to collect fictitious reviews. In other words, the authors had greater control in collecting fictitious reviews vis-à-vis authentic ones. However, this could not be

obviated because authentic reviews have been defined as those contributed by users who had paid and patronized a given hotel whereas fictitious reviews have been defined as those written by individuals based on imagination without any experience of staying in the hotel. The only feasible way to cater to this definition was to collect authentic reviews from authenticated websites and solicit fictitious reviews from participants in a research setting.

As a final remark, this paper reminds readers that the proposed theoretical model treated experience as a proxy for authenticity. In other words, it intended to distinguish between experiential and imaginative texts. Therefore, its validity should be viewed in this context without generalizing it to truths and lies. In any way, assessing truthfulness of user-generated content requires access to contributors' dispositions and motivations, which are seldom available.

It is perhaps high time all review websites allow submission of entries only after bona fide transactions. After all, prevention is better than cure. With such a mechanism in place, consumers would at least be assured that all reviews had been written by individuals who had post-purchase experience of the commodity that is being evaluated. Going forward, such a setting would shift the scholarly focus in the future from distinguishing between experiential and imaginative reviews toward the truth-lie dichotomy in the entries.

## **Appendix**

An additional analysis was conducted to examine if the proposed theoretical model could help distinguish between reviews posted in authenticated websites and those submitted in unauthenticated websites. The former was drawn from the 900 authentic reviews used in this paper while the latter included 800 reviews collected from TripAdvisor.com, an unauthenticated website, available from two publicly available datasets (Ott *et al.*, 2011; 2013). These reviews were analyzed using similar approaches used in the paper.

Since titles were not available in the publicly available datasets of TripAdvisor.com, only descriptions of reviews were analyzed. With descriptions alone, the model performed generally well in distinguishing between reviews from authenticated websites and those from unauthenticated websites (Omnibus test  $\chi^2 = 1254.15$ ,  $p < 0.001$ , Cox & Snell pseudo- $R^2 = 0.52$ , Nagelkerke Pseudo- $R^2 = 0.70$ , accuracy = 89.10%). This suggests that reviews from unauthenticated websites are not necessarily authentic, and could be assumed to be fictitious.

Reviews in authenticated websites were more difficult to be read ( $\text{Exp}(\beta)=1.55$ ), and were richer in nouns ( $\text{Exp}(\beta)=1.12$ ), exclusion words ( $\text{Exp}(\beta)=1.16$ ) as well as causal words ( $\text{Exp}(\beta)=1.23$ ) vis-à-vis reviews in unauthenticated websites. The former however contained fewer words ( $\text{Exp}(\beta)=0.98$ ), long words ( $\text{Exp}(\beta)=0.30$ ), articles ( $\text{Exp}(\beta)=0.82$ ), pronouns ( $\text{Exp}(\beta)=0.92$ ), singular self-references ( $\text{Exp}(\beta)=0.91$ ), firm words ( $\text{Exp}(\beta)=0.76$ ), and punctuations ( $\text{Exp}(\beta)=0.96$ ).

### **Acknowledgments:**

This work was supported by the Ministry of Education Research Grant AcRF Tier 2 (MOE2014-T2-2-020).

### **References**

- Al-Debei, M. M., Akroush, M. N., and Ashouri, M. I. (2015), "Consumer attitudes towards online shopping: The effects of trust, perceived benefits, and perceived web quality", *Internet Research*, Vol. 25 No. 5, pp. 707-733.
- Anderson, D.E., DePaulo, B.M., Ansfield, M.E., Tickle, J.J. and Green, E. (1999), "Beliefs about cues to deception: Mindless stereotypes or untapped wisdom?", *Journal of Nonverbal Behavior*, Vol. 23 No. 1, pp. 67-89.

- Banerjee, S. and Chua, A.Y.K. (2014), "Applauses in hotel reviews: Genuine or deceptive?", in *Science and Information Conference*, IEEE, New York, pp. 938-942.
- Bloomfield, R.J. (2002), "The 'incomplete revelation hypothesis' and financial reporting", *Accounting Horizons*, Vol. 16 No. 3, pp. 233-243.
- Boals, A. and Klein, K. (2005), "Word use in emotional narratives about failed romantic relationships and subsequent mental health", *Journal of Language and Social Psychology*, Vol. 24 No. 3, pp. 252-268.
- Buller, D.B. and Burgoon, J.K. (1996), "Interpersonal deception theory", *Communication Theory*, Vol. 6 No. 3, pp. 203-242.
- Cao, Q., Duan, W. and Gan, Q. (2011), "Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach", *Decision Support Systems*, Vol. 50 No. 2, pp. 511-521.
- Carbunar, B., Rahman, M., Pissinou, N., and Vasilakos, A. V. (2013), "A survey of privacy vulnerabilities and defenses in geosocial networks", *IEEE Communications Magazine*, Vol. 51 No. 11, pp. 114-119.
- Chall, J.S. and Dale, E. (1995), *Readability Revisited: The New Dale-Chall Readability Formula*, Brookline, Cambridge.
- 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.
- Chua, A.Y.K. and Banerjee, S. (2015), "Measuring the effectiveness of answers in Yahoo! Answers", *Online Information Review*, Vol. 39 No. 1, pp. 104-118.
- Crawford, M., Khoshgoftaar, T.M., Prusa, J.D., Richter, A.N. and Al Najada, H. (2015), "Survey of review spam detection using machine learning techniques", *Journal of Big Data*, Vol. 2 No. 1, pp. 1-24.

- Cohen, J. (1960), "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement*, Vol. 20 No. 1, pp. 37-46.
- Dawes, J. (2002), "Five point vs eleven point scales: Does it make a difference to data characteristics?", *Australasian Journal of Market Research*, Vol. 10 No. 1, pp. 39-47.
- DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K. and Cooper, H. (2003), "Cues to deception", *Psychological Bulletin*, Vol. 129 No. 1, pp. 74-118.
- Eisenhardt, K. (1989), "Building theories from case study research", *Academy of Management Journal*, Vol. 14 No. 4, pp. 532-550.
- Ekman, P. and Friesen, W.V. (1969), "Nonverbal leakage and clues to deception", *Psychiatry*, Vol. 32 No. 1, pp. 88-106.
- 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.
- Gunsch, M.A., Brownlow, S., Haynes, S.E. and Mabe, Z. (2000), "Differential linguistic content of various forms of political advertising", *Journal of Broadcasting & Electronic Media*, Vol 44 No. 1, pp. 27-42.
- Hancock, J.T., Curry, L.E., Goorha, S. and Woodworth, M. (2007), "On lying and being lied to: A linguistic analysis of deception in computer-mediated communication", *Discourse Processes*, Vol. 45 No. 1, pp. 1-23.
- Harris, C. (2012), "Detecting deceptive opinion spam using human computation", in *AAAI Workshop on Artificial Intelligence*, AAAI, Menlo Park, CA, pp. 87-93.
- Heydari, A., ali Tavakoli, M., Salim, N. and Heydari, Z. (2015), "Detection of review spam: A survey", *Expert Systems with Applications*, Vol. 42 No. 7, pp. 3634-3642.

- Hsu, M.H., Chuang, L.W. and Hsu, C.S. (2014), "Understanding online shopping intention: The roles of four types of trust and their antecedents", *Internet Research*, Vol. 24 No. 3, pp. 332-352.
- Humpherys, S.L., Moffitt, K.C., Burns, M.B., Burgoon, J.K. and Felix, W.F. (2011), "Identification of fraudulent financial statements using linguistic credibility analysis", *Decision Support Systems*, Vol. 50 No. 3, pp. 585-594.
- Jalilvand, M. R., and Samiei, N. (2012), "The impact of electronic word of mouth on a tourism destination choice: Testing the theory of planned behavior (TPB)", *Internet Research*, Vol. 22 No. 5, pp. 591-612.
- Jeong, H. J., and Koo, D. M. (2015), "Combined effects of valence and attributes of e-WOM on consumer judgment for message and product: The moderating effect of brand community type", *Internet Research*, Vol. 25 No. 1, pp. 2-29.
- Jindal, N. and Liu, B. (2008), "Opinion spam and analysis", in *Proceedings of the International Conference on Web search and Web Data Mining*, ACM, New York, pp. 219-230.
- Johnson, M.K. and Raye, C.L. (1981), "Reality monitoring", *Psychological Review*, Vol. 88 No. 1, pp. 67-85.
- Klein, D. and Manning, C.D. (2003), "Accurate unlexicalized parsing", in *Annual Meeting of the Association for Computational Linguistics*, ACL, Pennsylvania, pp. 423-430.
- Kugler, L. (2014), "Keeping online reviews honest", *Communications of the ACM*, Vol. 57 No. 11, pp. 20-23.
- Lee, K.Y. and Yang, S.B. (2015), "The role of online product reviews on information adoption of new product development professionals", *Internet Research*, Vol. 25 No. 3, pp. 435-452.

- Li, F., Huang, M., Yang, Y. and Zhu, X. (2011), "Learning to identify review spam", in *Proceedings of the International Joint Conference on Artificial Intelligence, AAAI*, California, pp. 2488-2493.
- Licht, M.H. (1995), "Multiple regression and correlation", in Grimm, L.G. and Yarnold, P.R. (Eds.), *Reading and Understanding Multivariate Statistics*, APA, Washington, pp. 19-64.
- Mackiewicz, J. (2015), "Quality in product reviews: What technical communicators should know", *Technical Communication*, Vol. 62 No. 1, pp. 3-18.
- McCornack, S.A. (1992), "Information manipulation theory", *Communications Monographs*, Vol. 59 No. 1, pp. 1-16.
- McLeod, J.M. and Pan, Z. (2005), "Concept explication and theory construction", in Dunwoody, S., Becker, L., McLeod, D. and Kosicki, G. (Eds.), *The Evolution of Key Mass Communication Concepts: Honoring Jack M. McLeod* (pp. 13-76). Creskill, NJ: Hampton Press.
- McQuaid, S.M., Woodworth, M., Hutton, E.L., Porter, S. and ten Brinke, L. (2015), "Automated insights: Verbal cues to deception in real-life high-stakes lies", *Psychology, Crime & Law*, Vol. 21 No. 7, pp. 617-631.
- Munar, A.M. and Jacobsen, J.K.S. (2014), "Motivations for sharing tourism experiences through social media", *Tourism Management*, Vol. 43, pp. 46-54.
- Newman, M.L., Pennebaker, J.W., Berry, D.S. and Richards, J.M. (2003), "Lying words: Predicting deception from linguistic styles", *Personality and Social Psychology Bulletin*, Vol. 29 No. 5, pp. 665-675.
- Ong, T., Mannino, M. and Gregg, D. (2014), "Linguistic characteristics of shill reviews", *Electronic Commerce Research and Applications*, Vol. 13 No. 2, pp. 69-78.

- Ott, M., Cardie, C. and Hancock, J.T. (2013), "Negative deceptive opinion spam", in *Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, ACL, Pennsylvania, pp. 497-501.
- Ott, M., Choi, Y., Cardie, C. and Hancock, J.T. (2011), "Finding deceptive opinion spam by any stretch of the imagination", in *Annual Meeting of the Association for Computational Linguistics*, ACL, Pennsylvania, pp. 309-319.
- Owusu, R. A., Mutshinda, C. M., Antai, I., Dadzie, K. Q., and Winston, E. M. (2016), "Which UGC features drive web purchase intent? A spike-and-slab Bayesian variable selection approach", *Internet Research*, Vol. 26 No. 1, pp. 22-37.
- Pan, M.C., Kuo, C.Y., Pan, C.T. and Tu, W. (2013), "Antecedent of purchase intention: Online seller reputation, product category and surcharge", *Internet Research*, Vol. 23 No. 4, pp. 507-522.
- Pasupathi, M. (2007), "Telling and the remembered self: Linguistic differences in memories for previously disclosed and previously undisclosed events", *Memory*, Vol. 15 No. 3, pp. 258-270.
- Pennebaker, J.W., Booth, R.J. and Francis, M.E. (2007), "Linguistic Inquiry and Word Count: LIWC [Computer software]", LIWC.net, Texas.
- Rayson, P., Wilson, A. and Leech, G. (2001), "Grammatical word class variation within the British National Corpus sampler", *Language and Computers*, Vol. 36 No. 1, pp. 295-306.
- Tausczik, Y.R. and Pennebaker, J.W. (2010), "The psychological meaning of words: LIWC and computerized text analysis methods", *Journal of Language and Social Psychology*, Vol. 29 No. 1, pp. 24-54.

- Weber, R. (2012), "Theory building in the information systems discipline: Some critical reflections" in Hart, D.N. and Gregor, S.D. (Eds.), *Information Systems Foundations: Theory Building in Information Systems*, ANU Press, Australia, pp. 1-20.
- Wu, G., Greene, D. and Cunningham, P. (2010), "Merging multiple criteria to identify suspicious reviews", in *Proceedings of the Conference on Recommender systems*, ACM, New York, NY, pp. 241-244.
- Yoo, K.H. and Gretzel, U. (2009), "Comparison of deceptive and truthful travel reviews", in Höpken, W., Gretzel, U. and Law, R. (Eds.), *Information and Communication Technologies in Tourism*, Springer, Vienna, pp. 37-47.
- Zhang, L. and Peng, T.Q. (2015), "Breadth, depth, and speed: Diffusion of advertising messages on microblogging sites", *Internet Research*, Vol. 25 No. 3, pp. 453-470.

**Table I.** Reviews received by the identified hotels as on 28 February, 2013

Tourist destinations	Hotel types	Names of the identified hotels	# Reviews			
			Agoda.com	Expedia.com	Hotels.com	Total
Bangkok	Luxury	lebua at State Tower	6,348	903	1,010	8,261
	Budget	Sawasdee Sukhumvit	951	9	15	975
	Mid-range	Nasa Vegas	5,802	142	268	6,212
Hong Kong	Luxury	L'Hotel Nina	6,603	155	369	7,127
	Budget	USA Hostel	3,074	71	165	3,310
	Mid-range	Casa Hotel	4,023	80	110	4,213
Kuala Lumpur	Luxury	Furama Bukit Bintang	5,973	134	263	6,370
	Budget	D'Oriental Inn	1,864	79	137	2,080
	Mid-range	Radius International	5,100	81	121	5,302
Singapore	Luxury	Mandarin Orchard	6,678	285	319	7,282
	Budget	Fragrance Ruby	2,300	34	47	2,381
	Mid-range	Ibis Bencoolen	4,077	258	302	4,637
Tokyo	Luxury	Metropolitan Tokyo	1,393	237	400	2,030
	Budget	Horidome Villa	741	107	309	1,157
	Mid-range	Shinjuku Washington	1,433	388	472	2,293

**Table II.** Indicators to distinguish between authentic and fictitious reviews

Constructs	Definitions	Operationalized Indicators	Related References
Comprehensibility	The extent to which a review is easy to understand.	(1) Mean readability, (2) familiar words, (3) characters per word, (4) words, (5) long words, (6) words per sentence	Cao <i>et al.</i> , 2011; Chall and Dale, 1995; Ermakova <i>et al.</i> , 2015; Ghose and Ipeirotis, 2011
Specificity	The extent to which a review is rich in details.	(1) Nouns, (2) adjectives, (3) prepositions, (4) articles, (5) conjunctions, (6) verbs, (7) adverbs, (8) pronouns, (9) singular self-references, (10) plural self-references, (11) visual words, (12) aural words, (13) feeling words, (14) temporal words, (15) spatial words, (16) lexical diversity, (17) function words	Hancock <i>et al.</i> , 2007; Ott <i>et al.</i> , 2011; Rayson <i>et al.</i> , 2001; Tausczik and Pennebaker, 2010; Yoo and Gretzel, 2009; Zhou <i>et al.</i> , 2003
Exaggeration	The writing style used in a review to sound convincing.	(1) Emotiveness, (2) positive words, (3) negative words, (4) past tense, (5) present tense, (6) future tense, (7) firm words, (8) upper case characters, (9) brand references, (10) ellipses, (11) exclamation points, (12) question marks, (13) all punctuations	Liu <i>et al.</i> , 2013; Newman <i>et al.</i> , 2003; Tausczik and Pennebaker, 2010; Yoo and Gretzel, 2009; Zhang and Peng, 2015; Zhou <i>et al.</i> , 2004
Negligence	Textual clues for deception detection that coincidentally leak out in a review.	(1) Discrepancy words, (2) filler words, (3) tentative words, (4) motion words, (5) exclusion words, (6) causal words, (7) insight words	Boals and Klein, 2005; McQuaid <i>et al.</i> , 2015; Newman <i>et al.</i> , 2003; Pasupathi, 2007; Tausczik and Pennebaker, 2010

**Table III.** Performance of the logistic regression models

Logistic Regression Results	Positive Reviews	Negative Reviews	Moderate Reviews
Omnibus test	$\chi^2 = 364.76^{***}$	$\chi^2 = 408.97^{***}$	$\chi^2 = 307.04^{***}$
Cox & Snell Pseudo-R <sup>2</sup>	0.45	0.49	0.40
Nagelkerke Pseudo-R <sup>2</sup>	0.61	0.66	0.53
Accuracy	82.17%	86.00%	78.83%

\*\*\* p < 0.001

**Table IV.** Odds ratios to differentiate between authentic and fictitious reviews

Measures	Positive Reviews		Negative Reviews		Moderate Reviews	
	Titles	Descs	Titles	Descs	Titles	Descs
Words	1.46 <sup>***</sup>	0.99	1.31 <sup>**</sup>	1.02 <sup>*</sup>	0.98	0.99
Nouns	1.01	1.03	1.00	1.16 <sup>*</sup>	1.04 <sup>***</sup>	1.18 <sup>*</sup>
Adjectives	1.02	0.85	0.99	0.87	1.02 <sup>*</sup>	0.70 <sup>*</sup>
Verbs	0.97	1.15	1.20 <sup>*</sup>	0.99	1.07	0.96
Adverbs	1.01	0.85	0.98	0.86	1.02	0.74 <sup>*</sup>
Singular self-references	1.02	0.78 <sup>***</sup>	1.04	0.75 <sup>***</sup>	0.92	0.99
Plural self-references	1.06	0.87 <sup>*</sup>	0.92	0.88	0.99	1.19 <sup>*</sup>
Visual words	1.00	0.94	0.92	1.08	1.05	0.72 <sup>**</sup>
Aural words	0.98	0.85	0.88 <sup>*</sup>	1.10	0.92	1.05
Feeling words	0.78	0.66 <sup>*</sup>	0.93	1.09	1.08	0.70 <sup>**</sup>
Spatial words	1.02	1.00	1.04 <sup>**</sup>	0.96	1.04 <sup>**</sup>	1.01
Temporal words	1.06	1.18 <sup>*</sup>	1.02	1.04	1.09 <sup>**</sup>	1.07
Positive emotion words	1.00	0.98	1.07 <sup>*</sup>	1.27 <sup>**</sup>	1.05 <sup>***</sup>	0.95
Negative emotion words	1.19	1.60 <sup>*</sup>	0.99	0.80 <sup>**</sup>	1.01	1.02
Future tense	1.12	1.64 <sup>*</sup>	0.67	1.52 <sup>*</sup>	0.85	1.35
Firm words	0.92 <sup>*</sup>	0.94	0.92	0.73 <sup>**</sup>	0.95	1.07
Brand References	0.75	0.89	0.46	0.35 <sup>**</sup>	0.55	0.80
Ellipses <sup>a</sup>	N/A	0.75	N/A	1.47 <sup>*</sup>	N/A	1.06
Exclamations	0.84 <sup>*</sup>	0.85	0.86 <sup>***</sup>	0.95	0.85 <sup>***</sup>	1.12
Discrepancy	0.85 <sup>*</sup>	0.79 <sup>*</sup>	0.94	1.01	0.89 <sup>*</sup>	1.18
Exclusion words	1.15 <sup>**</sup>	1.26 <sup>**</sup>	1.01	0.98	1.01	1.03
Causal words	1.09	1.02	1.02	1.35 <sup>**</sup>	0.99	0.99
Insight words	0.99	1.16	1.00	0.71 <sup>**</sup>	1.02	1.07

\*\*\* p < 0.001; \*\* p < 0.01; \* p < 0.05; <sup>a</sup> Ellipses did not occur at all among review titles. The corresponding odds ratios are therefore labelled as not applicable (N/A).

**Table V.** Summary of the differences between authentic and fictitious reviews

		Positive Reviews	Negative Reviews	Moderate Reviews
Comprehensibility	Titles	Titles of authentic reviews are longer than those of fictitious ones.	Titles of authentic reviews are longer than those of fictitious ones.	No differences
	Descs	No differences	Descriptions of authentic reviews are longer than those of fictitious ones.	No differences
Specificity	Titles	No differences	Titles of authentic reviews contain more verbs and spatial words but fewer aural words than those of fictitious ones.	Titles of authentic reviews contain more nouns, adjectives, spatial and temporal words than those of fictitious ones.
	Descs	Descriptions of authentic reviews contain more temporal words but fewer singular and plural self-references as well as feeling words than those of fictitious ones.	Descriptions of authentic reviews contain more nouns but fewer singular self-references than those of fictitious ones.	Descriptions of authentic reviews contain more nouns and plural self-references but fewer adjectives, adverbs, visual and feeling words than those of fictitious ones.
Exaggeration	Titles	Titles of authentic reviews contain fewer firm words and exclamations than those of fictitious ones.	Titles of authentic reviews contain more positive emotion words but fewer exclamations than those of fictitious ones.	Titles of authentic reviews contain more positive emotion words but fewer exclamations than those of fictitious ones.
	Descs	Descriptions of authentic reviews contain more negative emotion words and future tense than those of fictitious ones.	Descriptions of authentic reviews contain more positive emotion words, future tense and ellipses but fewer negative emotion words, firm words and brand references than those of fictitious ones.	No differences
Negligence	Titles	Titles of authentic reviews contain more exclusion words but fewer discrepancy words than those of fictitious ones.	No differences	Titles of authentic reviews contain fewer discrepancy words than those of fictitious ones.
	Descs	Descriptions of authentic reviews contain more exclusion words but fewer discrepancy words than those of fictitious ones.	Descriptions of authentic reviews contain more causal words but fewer insight words than those of fictitious ones.	No differences