| | |
|---|---|
| Title | Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels |
| Author(s) | Banerjee, Snehasish; Chua, Alton Yeow Kuan |
| Citation | Banerjee, S., & Chua, A. Y. K. (2016). Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels. Journal of Information Science, 43(1), 122-134. |
| Date | 2016 |
| URL | http://hdl.handle.net/10220/43563 |
| Rights | © 2016 The author(s) (published by SAGE Publications). This is the author created version of a work that has been peer reviewed and accepted for publication in Journal of Information Science, published by SAGE Publications on behalf of the author(s). 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/10.1177/0165551515625027]. |

*Article*

# Authentic versus Fictitious Online Reviews: A Textual Analysis across Luxury, Budget and Mid-Range Hotels

Journal of Information Science
1–13

DOI: 10.1177/0165551510000000
jis.sagepub.com
SAGE

**Snehasish Banerjee**
Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

**Alton Y. K. Chua**
Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

**Abstract**
Extant literature suggests that authentic and fictitious online reviews could be distinguished by leveraging on their textual characteristics. However, nuances in textual differences between authentic and fictitious reviews across different categories of hotels remain largely unknown. Therefore, this paper analyzes textual differences between authentic and fictitious reviews across three hotel categories, namely, luxury, budget and mid-range. It leverages on four possible textual characteristics—comprehensibility, specificity, exaggeration and negligence—that could offer clues to ascertain review authenticity. Using a dataset of 1,800 reviews (900 authentic + 900 fictitious), the results suggest that differences between authentic and fictitious reviews are largely inconsistent across hotel categories. This generally points to the difficulties in ascertaining review authenticity, which in turn offer implications for both research and practice.

**Keywords**
electronic word-of-mouth; online reviews; authenticity; text analysis; opinion spam

## 1. Introduction

With the rise of online communities, the volume of electronic word-of-mouth communication has increased dramatically. One of the popular genres of such communication takes the form of online hotel reviews, which are believed to report actual lodging experiences [1]. It is no surprise that potential travellers regularly turn to reviews to make booking decisions [2, 3, 4].

However, the Internet allows anyone—both bona fide and phoney alike—to submit reviews. Legitimate users can easily share their post-stay experience in hotels. It is equally easy for businesses to hire spammers to post fictitious reviews with the ulterior motive of gaining unfair edge over their rivals. Since fictitious reviews are written purposely to resemble authentic ones, users find it difficult to distinguish between the two [5, 6, 7].

To tackle this problem, scholars have mostly attempted to differentiate authentic from fictitious reviews by leveraging on their textual differences [6, 8, 9]. This is because texts written based on real experiences are likely to differ from those hinged on imagination [10, 11, 12]. Nevertheless, both theoretical and methodological research gaps persist.

On the theoretical front, extant literature lacks an overarching understanding of textual differences between authentic and fictitious reviews. For example, while [8] studied the use of superlatives in reviews, the level of details that offers clues to discern authenticity was ignored [13]. Conversely, while [14] studied the level of details in authentic and fictitious reviews, the resemblance of cover-up that offers tell-tale signs to infer authenticity was overlooked [12]. Hence, it is important to incorporate various aspects of textual characteristics to holistically identify the differences between authentic and fictitious reviews.

**Corresponding author:**
Snehasish Banerjee, 31 Nanyang Link, WKWSCI Building, Singapore 637718.
Email: snehasis002@ntu.edu.sg

On the methodological front, the textual characteristics used in prior studies to distinguish between authentic and fictitious reviews have not always been operationalized exhaustively. For example, [8] quantified comprehensibility of reviews as average word length without taking into account proxies such as text readability [15]. Likewise, [14] measured specificity by relying on the proportions of various parts-of-speech (POS) used in reviews but overlooked measures such as the use of spatial and temporal words [16]. Hence, the research gaps in extant literature call for a more exhaustive operationalization of textual characteristics to distinguish between authentic and fictitious reviews.

Additionally, related studies have mostly confined their datasets to reviews for popular hotels, and specifically, luxury properties. For example, [8] analyzed reviews for a Mariott hotel in the United States, while [6] studied entries for hotels in Chicago that were rated most highly in TripAdvisor. To glean nuanced perspectives, it is necessary to expand the scope of data collection to encompass various hotel categories such as luxury, budget and mid-range. This is because different hotel categories elicit different levels of expectation. In fact, the extent to which travellers' experience meets expectation is largely a function of hotel categories [17, 18, 19]. The writing style of reviews by both travellers as well as spammers could vary across luxury, budget and mid-range hotels.

Therefore, this paper seeks to predict authenticity of reviews using four textual characteristics, namely, comprehensibility, specificity, exaggeration and negligence. Comprehensibility refers to the reading ease of reviews. Specificity denotes the level of details in reviews. Exaggeration is a measure of hyperbolic tone used in reviews. Negligence points to indicators of cognitive load in reviews. Efforts were made to operationalize these textual characteristics as exhaustively as possible while avoiding redundancy. To delve deeper, the paper examines textual differences between authentic and fictitious reviews separately for luxury, budget as well as mid-range hotels. A total of 1,800 reviews (900 authentic + 900 fictitious) evenly distributed across the three hotel categories was used for investigation.

The rest of this paper proceeds as follows. The next section reviews the literature, which is followed by the research methods. The results are presented next followed by a discussion. The paper concludes by highlighting its contributions, limitations and future research directions.

## 2. Literature review

Prior studies that relied on textual measures to distinguish between authentic and fictitious information could be clustered into four major themes. These include comprehensibility, specificity, exaggeration and negligence. Studies on comprehensibility used textual measures such as "*readability*" [14: p. 502] and "*sentence length*" [20: p. 145]. Those related to specificity often conducted "*genre identification [informative or imaginative]*" [6: p. 310] or analyzed the use of "*details*" [21: p. 77]. Studies dealing with exaggeration used textual measures such as the use of "*sentiment [positive or negative]*" [22: p. 90] or "*punctuations*" [23: p. 465]. Finally, those pertaining to negligence tested the fraction of "*filler words*" [24: p. 190] or "*tentative constructions*" [21: p. 114]. However, these textual characteristics have seldom been conceptualized collectively to distinguish between authentic and fictitious reviews. They are further elaborated as follows.

### *2.1. Comprehensibility*

Comprehensibility refers to the extent to which reviews are clear and easy to understand [15]. Differences between authentic and fictitious reviews in terms of comprehensibility represent a double-edged sword. On the one hand, authentic reviews could be less comprehensible than fictitious ones because writing the former is cognitively easier [10, 25]. Individuals writing freely tend to use more sophisticated language than those writing under cognitively demanding conditions [26]. On the other hand, fictitious reviews could be less comprehensible than authentic ones because simplistic reviews are often considered incredulous [15]. Fictitious reviews might be purposely written using sophisticated language hoping that grandiloquence would lend credibility [8]. Therefore, the following research question (RQ) is presented:

- RQ 1: How do authentic reviews differ from fictitious ones in terms of comprehensibility across luxury, budget and mid-range hotels?

For the purpose of this paper, comprehensibility covers three sub-dimensions that include readability, word familiarity and structural features. Readability measures the effort and expertise required to grasp the meaning of reviews [27]. Word familiarity denotes the degree to which reviews use words that are easy to recognize [28]. Structural features refer to surface-level characteristics of reviews such as number of words, and number of long words [29].

Readable reviews with high word familiarity that are not overly verbose and those that use long words sparingly tend to enhance comprehensibility.

### 2.2. Specificity

Specificity refers to the extent of details indicated in reviews [13]. There are two competing views on the ways specificity of authentic and fictitious reviews might differ. The first holds that authentic reviews are more specific than fictitious ones because they are simply accounts of what has been experienced in reality [25]. The second view however suggests that fictitious reviews are more specific compared with authentic ones. This is because fictitious reviews, which are written without real experiences, could be over-compensated through concocted details [30]. Therefore, the following RQ is presented:

- RQ 2: How do authentic reviews differ from fictitious ones in terms of specificity across luxury, budget and mid-range hotels?

For the purpose of this paper, specificity covers four sub-dimensions that include informativeness, perceptual details, contextual details and lexical features. Informativeness refers to the richness of content in reviews [31]. Perceptual details indicate the degree to which reviews contain words that are used to convey a visceral experience [32]. Contextual details measure the extent to which reviews make spatial and temporal references [33]. Lexical features include word-level characteristics of reviews such as lexical diversity [34]. Informative reviews rich in perceptual as well as contextual details and those that are lexically rich tend to enhance specificity.

### 2.3. Exaggeration

Exaggeration refers to the ways specific types of words are used to construct sentences in reviews convincingly [35]. Differences between authentic and fictitious reviews in terms of exaggeration stem from the use of rhetorical strategies. Authentic reviews could resemble innocuous opinion-sharing entries written without intending to prove a point. In contrast, fictitious reviews could resemble attention-seeking entries written with an added conviction [8, 20, 36]. However, recent research suggests that fictitious reviews might be articulated with adequate guile to resemble authentic entries [14]. Therefore, the following RQ is presented:

- RQ 3: How do authentic reviews differ from fictitious ones in terms of exaggeration across luxury, budget and mid-range hotels?

For the purpose of this paper, exaggeration covers four sub-dimensions that include affective cues, tenses, emphases and punctuations. Affective cues refer to the use of emotion words that are used to create a lasting impression [36]. Tenses indicate the temporal focus of reviews [16]. A higher proportion of present or future tense than past tense in reviews suggests experiences that are likely to persist now or re-appear later [37]. Emphases through the use of rhetorical devices such as upper-case characters denote the hyperbolic nature of reviews [8]. Punctuations such as question marks and exclamation marks also tend to convey exaggeration [23].

### 2.4. Negligence

Negligence refers to cues that leak out due to the cognitive challenges involved in writing fictitious reviews [38]. Writing fictitious reviews is cognitively more challenging than writing authentic entries [10, 25]. Individuals writing fictitious reviews could get aroused psychologically. The heightened arousal, which is difficult to mask [39], could lead to the emanation of negligence cues in two ways. One, it might prevent individuals from writing fictitious reviews sincerely. Alternatively, it could sub-consciously entice them to go overboard in an attempt to cover-up. Both of these result in the leakage of cues which could be detected [12]. Therefore, the following RQ is presented:

- RQ 4: How do authentic reviews differ from fictitious ones in terms of negligence across luxury, budget and mid-range hotels?

For the purpose of this paper, negligence covers three sub-dimensions that include self-references, uncertainty cues and cognitive cues. Self-references reflect the degree of psychological immediacy expressed in reviews. They could be used less in fictitious reviews than authentic ones to reduce accountability [16]. Uncertainty cues connote a lack of conviction in reviews [40]. Cognitive cues refer to the degree of mental processing involved in writing reviews [16]. The four textual characteristics to distinguish between authentic and fictitious reviews, namely, comprehensibility, specificity, exaggeration and negligence, along with their sub-dimensions are presented in Table 1.

**Table 1.** The four textual characteristics to distinguish between authentic and fictitious reviews.

| Textual characteristics | Sub-dimensions | References |
|---|---|---|
| Comprehensibility | Readability | [27] |
| | Word familiarity | [28] |
| | Structural features | [29] |
| Specificity | Informativeness | [31] |
| | Perceptual details | [32] |
| | Contextual details | [33] |
| | Lexical features | [34] |
| Exaggeration | Affective cues | [36] |
| | Tenses | [16] |
| | Emphases | [8] |
| | Punctuations | [23] |
| Negligence | Self-references | [16] |
| | Uncertainty cues | [40] |
| | Cognitive cues | [16] |

## 3. Research Methods

### *3.1. Data sources for authentic Reviews*

Identification of data sources for authentic reviews involved two steps: selecting authenticated review websites, and selecting hotels. The first step required selecting websites that only accept entries from bona fide travellers. Such entries are guaranteed to have been written after post-stay experience in hotels. For this purpose, the top 10 travel websites based on Alexa's user traffic statistics[1] were short-listed for selection on February 28, 2013. These included Agoda.com, Booking.com, Expedia.com, Hotels.com, Kayak.com, Priceline.com, Southwest.com, TripAdvisor.com, United.com and Xe.com. Among these, six authenticated review websites included Agoda.com, Booking.com, Expedia.com, Hotels.com, Kayak.com and Priceline.com. However, not all offered comparable functionalities. While some solicited reviews as a combination of titles and descriptions, others sought only descriptions. To account for greater textual nuances, this paper relied on reviews containing titles as well as descriptions since both components play crucial communicative roles in reviews [41]. Hence, review websites that do not have the provision for submitting titles, namely, Booking.com, Kayak.com and Priceline.com, were excluded. The remaining three review websites, namely, Agoda.com, Expedia.com and Hotels.com, were selected.

The second step required selecting hotels that attract large volumes of authentic reviews. This would ensure an adequate number of entries from which authentic reviews could be admitted into the dataset. Hotels located in five popular tourist destinations within Asia—Bangkok, Hong Kong, Kuala Lumpur, Singapore and Tokyo—were chosen. Specifically, from each tourist destination, three properties that attract numerous reviews in the selected websites were randomly identified to yield a list of 15 hotels (5 tourist destinations x 3 hotels). Each set of three properties comprised a luxury hotel, a budget hotel, and a mid-range hotel. Thus, the identified list of hotels comprised five luxury hotels, five budget hotels, and five mid-range hotels. All hotels had attracted more than 1,000 reviews cumulatively across the selected websites. The validity of categorizing hotels as luxury, budget or mid-range was ensured through a two-pronged strategy. One, the hotels had to feature in all the three selected websites. Two, their hotel categories had to match consistently across the three websites. Four or five star-rated properties were deemed as luxury hotels, one or two star-rated properties were deemed as budget hotels, and three star-rated properties were deemed as mid-range hotels.

### *3.2. Collection of authentic reviews*

For each of the 15 selected hotels (5 luxury + 5 budget + 5 mid-range), a total of 60 authentic reviews (20 positive + 20 negative + 20 moderate) were collected to yield 900 entries altogether (15 hotels x 60 reviews). The reviews uniformly straddled across the three sentiments—positive, negative and moderate—to enhance generalizability of the study [7]. Sentiments expressed in reviews are often found correlated with ratings albeit weakly [42, 43]. Hence, informed by the approaches used in prior studies [44, 45], review sentiments were ascertained based on the polarity of review ratings.

Specifically, Expedia.com and Hotels.com use a 5-point rating scale, whereas Agoda.com employs a 10-point rating scale. Scales that differ in their ranges cannot be linearly interpolated [46]. In other words, a rating of 1 on a 5-point scale is not necessarily 2 on a 10-point scale [47]. To make ratings from the three websites comparable to one another, the rescaling approach of [48] was utilized in ascertaining review sentiments.

Authentic reviews were admitted into the dataset based on four criteria. First, they had to be posted as recently as possible. Second, both titles and descriptions of reviews had to contain meaningful English text. Third, descriptions of reviews had to be at least more than 150 characters in length to allow for a meaningful analysis [6]. Fourth, the reviews' contributors had to disclose their country of origin.

For each review, data from the following fields were collected: numerical rating, title, description, and contributors' country of origin. The 900 authentic reviews were contributed by bona fide travellers from across the four major geographical regions of the world in the following proportions: 71 from America, 730 from Asia Pacific, 88 from Europe, and 11 from Middle East as well as Africa. It was necessary to track this so that similar proportions could also be created in the corpus of fictitious reviews.

### 3.3. Collection of fictitious reviews

Fictitious reviews were solicited from participants who had no prior experience of staying in the selected hotels. Over 400 participants were recruited via convenient sampling and snowballing while keeping in mind the proportions of authentic reviews obtained from the four major geographical regions. Adequate efforts were made to ensure that the proportions of contributors' country of origin in the fictitious reviews was comparable to those in the authentic entries. This would help to control for cultural differences in writing style, thus facilitating a fair investigation.

All participants aged 21 to 45 years, and were either undergraduate students or possessed graduate degrees. After all, reviews are mostly written by young and educated individuals [49]. Moreover, they were regular readers or contributors of review websites, and had travel experience during the last year.

The instruction given to participants for writing fictitious reviews was informed by prior studies [6, 8]. Participants could choose to write at most six realistic but fictitious reviews for six different hotels, whose website URLs were provided. Some participants were instructed to write positive reviews, some were asked to submit negative reviews, while the rest were required to contribute moderate entries. All reviews had to be in English, with meaningful titles and meaningful descriptions of at least 150 characters.

After collecting fictitious reviews for a period of more than six months, a total of some 950 reviews were obtained. All the reviews were manually inspected to ensure that they were meaningful. Of these, 900 fictitious reviews (15 hotels x 60 reviews) evenly spread across the 15 identified hotels as well as positive, negative and moderate sentiments were randomly admitted into the dataset by ensuring maximal comparability with the authentic entries in terms of contributors' country of origin. Specifically, based on participants' country of origin, the 900 fictitious reviews submitted from the four major geographical regions of the world were in the following proportions: 69 from America, 732 from Asia Pacific, 85 from Europe, and 14 from Middle East and Africa. The corpus of authentic reviews and that of fictitious reviews contained comparable proportions of entries submitted by individuals from the four geographical regions. This helped to control for differences in writing style—that might stem from cultural differences—as much as possible, thereby affording a fair comparison between the two corpora.

### 3.4. Measures of the textual characteristics

With respect to comprehensibility, readability is commonly measured using indicators such as Automated-Readability Index, Coleman-Liau Index, Flesch-Kincaid Grade Level, Gunning-Fog Index, Lasbarhets Index, and Rate Index [15, 50, 51, 52]. Lower values of the indicators suggest more readable reviews. To incorporate the merits of the six indicators [15], they were averaged to create a composite variable (henceforth, mean readability index). For word familiarity, each word in reviews was compared against the Dale-Chall lexicon of familiar words to measure the proportion of easily-recognizable words [28]. Structural features were measured as the number of characters per word, number of words, fraction of words with 10 or more characters (henceforth, long words), and number of words per sentence [8, 20, 22, 29].

With respect to specificity, informativeness was measured based on the proportion of the following eight POS tags: nouns, adjectives, prepositions, articles, conjunctions, verbs, adverbs, and pronouns. In informative texts, the first four are generally abundant, and the next four generally scanty [6, 16, 31]. Perceptual details were measured as the proportion of words in reviews that connote visual (e.g., view), aural (e.g., listen), and feeling (e.g., touch) perceptions

[10, 33]. Contextual details were measured based on the fraction of spatial (e.g., down), and temporal (e.g., until) words [12, 53]. Lexical features entailed lexical diversity, emotiveness, and proportion of function words [8, 16, 54].

With respect to exaggeration, the use of affective words was measured as the fraction of positive emotion (e.g., nice), and negative emotion (e.g., ugly) words [8, 55]. Tenses were measured as the fraction of past, present and future tense words used in reviews [16, 37]. Emphases were operationalized in terms of the proportion of firm words (e.g., always), fraction of upper case characters, and references to the hotel name (henceforth, brand references) [8, 16]. Punctuations were measured in terms of the use of ellipses, emoticons, exclamation marks, question marks, and all punctuations used in reviews [20, 23, 56, 57].

With respect to negligence, self-references were measured as the fraction of first person singular (e.g., I), and plural (e.g., we) words [30]. Uncertainty cues were measured as the proportion of modal verbs (e.g., could), filler words (e.g., you know), and tentative words (e.g., perhaps) [54, 58]. Cognitive cues were operationalized based on the fraction of causal (e.g., because), insight (e.g., think), motion (e.g., arrive), and exclusion (e.g., without) words [25, 59, 60].

As shown in Table 2, the textual characteristics were measured using a total of 44 variables (6 for comprehensibility + 16 for specificity + 13 for exaggeration + 9 for negligence). These were calculated separately for titles and descriptions of all the 1,800 reviews in the dataset (900 authentic + 900 fictitious) using the Linguistic Inquiry and Word Count (LIWC) tool [61], Stanford Parser's POS tagger [62], as well as some custom-developed Java programs.

### *3.5. Data analysis*

The statistical procedure of logistic regression was used for data analysis. It is appropriate when a problem involves a dichotomous dependent variable, and a set of either continuous or categorical predictor variables. It allows assessing the extent to which the set of predictor variables explains or predicts the categorical dependent variable [63, 64, 65].

In this paper, the dependent variable was review authenticity, which was dichotomous in nature (1 = authentic, 0 = fictitious). The textual measures of comprehensibility, specificity, exaggeration and negligence were the predictor variables, all of which were continuous. Therefore, using logistic regression was appropriate. Specifically, it modelled the log odds of a review being authentic versus it being fictitious [63, 66].

For review titles, only 40 of the 44 variables (cf. Table 2) were used as predictors. Mean readability (variable #1), and number of words per sentence (variable #6) depend on sentence counts. However, review titles rarely contain sentences. Moreover, ellipses (variable #31), and emoticons (variable #32) in review titles yielded few occurrences in the dataset. On the other hand, for review descriptions, 43 of the 44 variables were used as predictors. The use of emoticons (variable #32) was excluded as it yielded few occurrences in the dataset.

Thus, the logistic regression analysis used a total of 83 textual measures (40 for titles + 43 for descriptions) to predict the dichotomous dependent variable review authenticity. The analysis was repeated thrice for luxury, budget and mid-range hotels. All analyses were conducted using the SPSS (v21) software. A p-value of 0.05 or less was used as the threshold for statistical significance. The overall performance of the model was checked using the $\chi^2$ statistic, statistical significance, deviance and pseudo-$R^2$ values. To delve deeper, the relationship between each predictor variable and review authenticity was checked in terms of odds ratio [64, 67]. Such an approach is widely documented in prior studies [66, 68, 69].

Logistic regression is however sensitive to overly high correlations among predictor variables—a condition known as multicollinearity [64]. To check for possible multicollinearity, the pair-wise inter-correlations among the predictor variables were examined prior to the analyses. All values were less than 0.80, suggesting that multicollinearity was not a concern [70, 71].

## 4. Results

The logistic regression model demonstrated promising results in distinguishing between authentic and fictitious reviews for luxury ($\chi^2$ = 309.27, $p < 0.001$, Deviance = 522.51, Pseudo-$R^2$ = 53.70%), budget ($\chi^2$ = 346.64, $p < 0.001$, Deviance = 485.14, Pseudo-$R^2$ = 58.50%), as well as mid-range hotels ($\chi^2$ = 295.57, $p < 0.001$, Deviance = 535.21, Pseudo-$R^2$ = 52.00%). The model accounting for over 50% variability in the outcome variable points to its robustness. After all, research on user-generated content often yields much lower $R^2$ values [72, 73].

The likelihood ratio tests for the predictors with statistically significant odds ratio—Exp(β)—in distinguishing between authentic and fictitious reviews across luxury, budget and mid-range hotels are presented below (cf. Table 3). For luxury hotels, titles of authentic reviews differed significantly from those of fictitious ones in terms of four predictors. Specifically, the former was richer in nouns (Exp(β) = 1.02, $p < 0.05$), but contained fewer firm words

(Exp(β) = 0.90, $p < 0.01$), brand references (Exp(β) = 0.26, $p < 0.05$), and exclamation marks (Exp(β) = 0.93, $p < 0.001$). Besides, descriptions of authentic reviews differed significantly from those of fictitious ones in terms of five predictors. In particular, the former was richer in future tense (Exp(β) = 1.92, $p < 0.01$), but contained fewer pronouns (Exp(β) = 0.90, $p < 0.05$), visual words (Exp(β) = 0.80, $p < 0.05$), firm words (Exp(β) = 0.85, $p < 0.05$), and first person singular words (Exp(β) = 0.85, $p < 0.05$).

**Table 2.** Measures of the textual characteristics.

| Textual characteristics | Sub-dimensions | Variables | References |
|---|---|---|---|
| Comprehensibility | Readability | (1) Mean readability index | [15] |
| | Word familiarity | (2) Dale-Chall words | [28] |
| | Structural features | (3) Characters per word | [8, 20, 22, 29] |
| | | (4) Words | |
| | | (5) Long words | |
| | | (6) Words per sentence | |
| Specificity | Informativeness | (7) Nouns | [6, 16, 31] |
| | | (8) Adjectives | |
| | | (9) Prepositions | |
| | | (10) Articles | |
| | | (11) Conjunctions | |
| | | (12) Verbs | |
| | | (13) Adverbs | |
| | | (14) Pronouns | |
| | Perceptual details | (15) Visual perceptions | [10, 33] |
| | | (16) Aural perceptions | |
| | | (17) Feeling perceptions | |
| | Contextual details | (18) Spatial words | [12, 53] |
| | | (19) Temporal words | |
| | Lexical features | (20) Lexical diversity | [8, 16, 54] |
| | | (21) Emotiveness | |
| | | (22) Function words | |
| Exaggeration | Affective cues | (23) Positive emotion words | [8, 55] |
| | | (24) Negative emotion words | |
| | Tenses | (25) Past tense | [16, 37] |
| | | (26) Present tense | |
| | | (27) Future tense | |
| | Emphases | (28) Firm words | [8, 16] |
| | | (29) Upper case characters | |
| | | (30) Brand References | |
| | Punctuations | (31) Ellipses | [20, 23, 56, 57] |
| | | (32) Emoticons | |
| | | (33) Exclamation marks | |
| | | (34) Question marks | |
| | | (35) All punctuations | |
| Negligence | Self-references | (36) First person singular words | [30] |
| | | (37) First person plural words | |
| | Uncertainty cues | (38) Modal verbs | [54, 58] |
| | | (39) Filler words | |
| | | (40) Tentative words | |
| | Cognitive cues | (41) Causal words | [25, 59, 60] |
| | | (42) Insight words | |
| | | (43) Motion words | |
| | | (44) Exclusion words | |

For budget hotels, titles of authentic reviews differed significantly from those of fictitious ones in terms of 13 predictors. Specifically, the former was lengthier (Exp(β) = 1.19, $p < 0.05$) with more nouns (Exp(β) = 1.02, $p < 0.05$), prepositions (Exp(β) = 1.08, $p < 0.05$), adverbs (Exp(β) = 1.07, $p < 0.05$), spatial words (Exp(β) = 1.03, $p < 0.05$), temporal words (Exp(β) = 1.05, $p < 0.05$), positive emotion words (Exp(β) = 1.03, $p < 0.05$), causal words (Exp(β) = 1.22, $p < 0.01$), and exclusion words (Exp(β) = 1.08, $p < 0.01$), but with fewer function words (Exp(β) = 0.94, $p < 0.05$), exclamation marks (Exp(β) = 0.85, $p < 0.001$), modal verbs (Exp(β) = 0.83, $p < 0.01$), and insight words (Exp(β) = 0.90, $p < 0.05$). Besides, descriptions of authentic reviews differed significantly from those of fictitious ones in terms of nine predictors. In particular, the former was richer in temporal words (Exp(β) = 1.10, $p < 0.05$), causal words (Exp(β) = 1.16, $p < 0.05$), and exclusion words (Exp(β) = 1.13, $p < 0.05$), but contained fewer articles (Exp(β) = 0.82, $p < 0.001$), visual words (Exp(β) = 0.73, $p < 0.05$), brand references (Exp(β) = 0.50, $p < 0.05$), first person singular words (Exp(β) = 0.72, $p < 0.001$), first person plural words (Exp(β) = 0.86, $p < 0.05$), and tentative words (Exp(β) = 0.82, $p < 0.05$).

For mid-range hotels, titles of authentic reviews differed significantly from those of fictitious ones in terms of seven predictors. Specifically, the former was lengthier (Exp(β) = 1.32, $p < 0.01$) with more nouns (Exp(β) = 1.02, $p < 0.01$), spatial words (Exp(β) = 1.05, $p < 0.001$), temporal words (Exp(β) = 1.07, $p < 0.01$), and positive emotion words (Exp(β) = 1.04, $p < 0.01$), but with fewer brand references (Exp(β) = 0.44, $p < 0.01$), and exclamation marks (Exp(β) = 0.90, $p < 0.001$). Besides, descriptions of authentic reviews differed significantly from those of fictitious ones in terms of six predictors. In particular, the former was richer in verbs (Exp(β) = 1.28, $p < 0.05$), and exclusion words (Exp(β) = 1.14, $p < 0.05$), but contained fewer adjectives (Exp(β) = 0.75, $p < 0.05$), negative emotion words (Exp(β) = 0.80, $p < 0.01$), past tense (Exp(β) = 0.80, $p < 0.05$), and brand references (Exp(β) = 0.43, $p < 0.01$).

Interestingly, budget hotels showed the most number of significant variables in distinguishing between authentic and fictitious reviews (13 + 9 = 22), followed by mid-range hotels (7 + 6 = 13), and luxury hotels (4 + 5 = 9). This suggests that reviews for budget hotels offer most opportunities to detect differences between authentic and fictitious entries, whereas those for luxury hotels lie at the opposite end of the spectrum. Furthermore, statistically significant predictor variables corresponding to titles (13 + 7 + 4 = 24) outnumbered those related to descriptions (9 + 6 + 5 = 20) of reviews across all hotel categories. Review titles probably offer more chances to detect differences between authentic and fictitious entries compared with review descriptions.

## 5. Discussion

The paper gleans four findings from Table 3 corresponding to the four textual characteristics, each discussed separately in terms of titles and descriptions. First, in terms of comprehensibility, titles of authentic reviews for budget and mid-range hotels were lengthier than their fictitious counterparts. A possible explanation is that individuals writing fictitious text often refrain from providing verbose entries so that little cues for detection are emanated [74]. Therefore, titles of these fictitious reviews were perhaps short. Such a difference in titles between authentic and fictitious reviews was however not observed for luxury hotels.

Differences in comprehensibility of review descriptions between authentic and fictitious entries were blurred regardless of hotel categories. Prior studies suggested mixed possibilities on how comprehensibility of authentic and fictitious reviews could differ [15, 25, 26]. While some posited authentic reviews to be more comprehensible, others expected the converse to be true. However, to the best of the authors' knowledge, this study is one of the earliest to empirically analyze comprehensibility using a wide range of measures that includes multiple readability indicators, word familiarity, and several structural features. The findings suggest that comprehensibility of review descriptions might not be a useful textual characteristic to distinguish between authentic and fictitious reviews.

Second, in terms of specificity, titles of authentic reviews were richer in nouns than those of fictitious entries regardless of hotel categories. Nouns rank among the most content-rich word categories [34], which were expectedly abundant in the former. After all, authentic reviews written after real experiences could be more content-rich than fictitious entries [10, 12]. For budget and mid-range hotels, titles of authentic reviews were richer in spatial as well as temporal words compared with those of fictitious entries. This rendered the former more contextually detailed [33]. Authentic reviews, which are supposedly informative, could be rich in prepositions with limited function words [31]. This was evident in titles of authentic reviews only for budget hotels. Contrary to literature however [16, 31], they were found to contain more adverbs than titles of fictitious reviews.

Differences in specificity of review descriptions between authentic and fictitious entries were largely inconsistent across the three hotel categories. For luxury hotels, authentic reviews were scantier in pronouns and visual words vis-à-vis fictitious ones. For budget hotels, authentic reviews contained more temporal words but fewer articles and visual words than fictitious ones. For mid-range hotels, authentic reviews contained more verbs but fewer adjectives compared with fictitious ones. Among POS tags, prior research suggests that authentic reviews might be richer in nouns, adjectives, prepositions, and articles yet containing fewer conjunctions, verbs, adverbs, and pronouns compared with fictitious entries [6, 16, 31, 55]. However, such differences were largely blurred. Moreover, perceptual details in the form of visual words seem to be easily concocted in fictitious reviews to resemble authentic ones.

**Table 3.** Odds ratio for the predictors to distinguish between authentic and fictitious reviews.

| Measures | Luxury hotels | | Budget hotels | | Mid-range hotels | |
|---|---|---|---|---|---|---|
| | Titles | Descriptions | Titles | Descriptions | Titles | Descriptions |
| **Comprehensibility** | | | | | | |
| Words | 1.12 | 1.00 | 1.19* | 1.00 | 1.32** | 1.00 |
| **Specificity** | | | | | | |
| Nouns | 1.02* | 1.06 | 1.02* | 1.10 | 1.02** | 1.07 |
| Adjectives | 1.00 | 0.83 | 1.00 | 0.88 | 1.01 | 0.75* |
| Prepositions | 0.98 | 0.98 | 1.08* | 0.99 | 0.99 | 1.02 |
| Articles | 1.00 | 0.90 | 1.01 | 0.82*** | 0.99 | 0.94 |
| Verbs | 1.07 | 0.96 | 0.98 | 0.99 | 0.84 | 1.28* |
| Adverbs | 1.01 | 0.88 | 1.07* | 0.83 | 0.98 | 0.83 |
| Pronouns | 1.00 | 0.90* | 1.03 | 0.95 | 1.01 | 0.94 |
| Visual words | 0.98 | 0.80* | 0.99 | 0.73* | 1.10 | 1.11 |
| Spatial words | 1.01 | 1.02 | 1.03* | 1.01 | 1.05*** | 0.99 |
| Temporal words | 1.03 | 0.99 | 1.05* | 1.10* | 1.07** | 1.06 |
| Function words | 1.00 | 0.95 | 0.94* | 1.07 | 1.01 | 0.99 |
| **Exaggeration** | | | | | | |
| Positive emotion words | 1.02 | 0.99 | 1.03* | 1.03 | 1.04** | 0.97 |
| Negative emotion words | 1.01 | 1.04 | 0.99 | 1.06 | 0.97 | 0.80** |
| Past tense | 0.94 | 1.05 | 1.03 | 0.96 | 1.19 | 0.80* |
| Future tense | 0.95 | 1.92** | 1.09 | 1.29 | 1.19 | 0.93 |
| Firm words | 0.90** | 0.85* | 0.98 | 0.93 | 0.96 | 0.96 |
| Brand references | 0.26* | 0.93 | 1.64 | 0.50* | 0.44** | 0.43** |
| Exclamation marks | 0.93*** | 1.06 | 0.85*** | 0.87 | 0.90*** | 0.96 |
| **Negligence** | | | | | | |
| First person singular words | 1.03 | 0.85* | 1.06 | 0.72*** | 0.94 | 0.91 |
| First person plural words | 0.98 | 0.99 | 0.89 | 0.86* | 0.21 | 0.89 |
| Modal verbs | 0.97 | 1.05 | 0.83** | 1.03 | 0.94 | 0.88 |
| Tentative words | 0.98 | 0.89 | 0.95 | 0.82** | 0.98 | 0.95 |
| Causal words | 1.04 | 1.07 | 1.22** | 1.16* | 1.00 | 1.14 |
| Insight words | 1.03 | 1.07 | 0.90* | 0.98 | 1.01 | 0.99 |
| Exclusion words | 1.02 | 0.99 | 1.08** | 1.13* | 1.04 | 1.14* |

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Third, in terms of exaggeration, titles of authentic reviews contained fewer exclamation marks than those of fictitious ones regardless of hotel categories. As suggested in prior research [23], punctuations seem to offer tell-tale signs to predict authenticity. Additionally, brand references were scanty in titles of authentic reviews for luxury and mid-range hotels. Those for budget and mid-range hotels were replete with positive emotion words. Besides, consistent with expectation [16], titles of authentic reviews for luxury hotels contained fewer firm words compared with those of fictitious entries.

Review descriptions of authentic reviews contained fewer brand references vis-à-vis their fictitious counterparts for budget and mid-range hotels. Although [8] obtained such a finding for luxury hotels, this paper does not find a similar trend. Apart from the use of brand references, no other measure of exaggeration could differentiate between descriptions of authentic and fictitious reviews for budget hotels. Nonetheless, authentic reviews for luxury hotels contained more future tense words and fewer firm words, whereas those for mid-range hotels included fewer negative emotion words and past tense compared with their fictitious counterparts.

Fourth, in terms of negligence, titles of authentic and fictitious reviews differed only for budget hotels. Consistent with the literature [16, 25, 29], authentic reviews contained more exclusion words but fewer modal verbs and insight words vis-à-vis fictitious entries. However, contrary to prior studies [16], authentic reviews contained more causal words than fictitious entries did. This again serves as growing evidence that it is hardly challenging to write fictitious reviews that resemble authentic ones [7, 14].

Differences in negligence of review descriptions between authentic and fictitious entries were largely inconsistent across the three hotel categories. For luxury hotels, authentic reviews contained fewer first person singular words compared with fictitious entries. For budget hotels, authentic reviews included more causal and exclusion words but with fewer first person singular as well as plural words, and tentative words vis-à-vis fictitious ones. For mid-range hotels, authentic reviews contained more exclusion words than fictitious entries. The result that authentic reviews were fraught with causal words was an aberration from prior research [16]. Even though authentic reviews are expected to contain fewer cognitive cues than fictitious entries, the former could still be richer in causal words. This suggests that authentic reviews might not always reflect the ideal properties of authentic texts. Herein lies the conundrum of reviews in general. While fictitious reviews are usually written deliberately to pass off as authentic, some authentic reviews could be mistakenly thought of as being fictitious.

## 6. Conclusion

This paper analyzed textual differences between authentic and fictitious reviews across luxury, budget as well as mid-range hotels. Four textual characteristics were considered, namely, comprehensibility, specificity, exaggeration and negligence. The differences between authentic and fictitious reviews emerged as being largely inconsistent across hotel categories.

The paper contributes in three ways. First, it represents one of the earliest attempts to analyze differences between authentic and fictitious reviews—comprising titles and descriptions—separately across luxury, budget as well as mid-range hotels. The motivation for such an analysis was based on the premise that the differences might not necessarily be consistent across hotel categories [17]. The results generally lend support for the premise, thereby confirming the merit of the approach. Additionally, the paper dovetails prior studies [8, 14] by presenting a comprehensive set of measures to operationalize comprehensibility, specificity, exaggeration and negligence.

Second, this paper finds that the number of variables that significantly distinguished between authentic and fictitious reviews was the least for luxury hotels among the hotel categories. Interestingly, related prior studies have mostly looked into reviews for luxury hotels [6, 7, 8, 55]. This creepy coincidence raises question on whether scholars, by publishing their results, are making it easier for spammers to write authentic-like fictitious reviews. It is certainly plausible to learn from such results, and tweak strategies of writing fictitious reviews in order to pass them off as authentic. It is unfortunate that even though scholars had been grappling with the problem of review authenticity, their efforts apparently have a huge potential to boomerang. As fictitious reviews could be written by crafting new devious strategies to game the system, research always seems to be playing the catch-up.

Third, this paper exposes the conundrum of reviews on the Internet. Despite being widely used for purchase decision-making [2, 3, 4], reviews need to be taken with a pinch of salt. This is because of the growing prevalence of authentic-like fictitious reviews, which are often the by-products of paid marketing lies. Furthermore, the differences between authentic and fictitious reviews could be disparate across various types of products or services. This makes it almost impossible for both computational algorithms as wells as humans to devise standardized fool-proof strategies to discern review authenticity. Perhaps, there exists no easy solution to address this stubborn problem. Nonetheless, a baby step involves websites allowing submission of reviews only after bona fide transactions. Even then, it could still be a challenge to ensure that such entries honestly describe truthful experiences. In the long run, regulators need to seriously consider developing stringent cyber laws and industry codes to curb the extent to which businesses could engage in review fakery [75].

There are four major limitations in this paper. First, it analyzed review authenticity for hotels located in five tourist destinations within Asia. Caution needs to be exercised in generalizing the findings to reviews for hotels located in other tourist destinations. Second, even though participants were asked to write fictitious reviews with instructions informed by prior studies [6, 8], the extent to which they were motivated and capable to do the task in the first place could not be verified. Third, this paper conducted the textual analysis of authentic and fictitious reviews without taking noise in the texts into account. In any way, most analysis such as POS tagging is known to be "*relatively robust regarding noisy texts*" [76: p. 904]. Fourth, this paper textually analyzed authentic and fictitious reviews that were written only in English. Hence, these findings are not intended to be applied to non-English entries.

Nonetheless, this paper offers new research directions. One direction involves investigating the extent to which textual characteristics of reviews flag off entries as suspicious across different types of websites such as TripAdvisor, which does not require authentication, and Expedia, which solicits reviews only from bona fide travellers. Such studies could help identify websites with a relatively superior working model in thwarting fictitious reviews. Another research direction includes analyzing writing strategies of fictitious reviews across hotel categories. Although much has been done on ways to predict authenticity, little efforts have hitherto been invested in studying the process of writing fictitious reviews. Similar studies could also be replicated for non-English reviews in evaluation of other products and services to further extend the understanding of authenticity in the domain of electronic word-of-mouth in general.

## Notes

1. http://www.alexa.com/topsites/category/Top/Recreation/Travel

## Acknowledgements

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

## References

[1] Martínez-Cámara E, Martín-Valdivia MT, Ureña-López LA and Mitkov R. Polarity classification for Spanish tweets using the COST corpus. *Journal of Information Science* 2015; 41(3): 263-272.

[2] Brilhante IR, Macedo JA, Nardini FM, Perego R and Renso C. On planning sightseeing tours with TripBuilder. *Information Processing & Management* 2015; 51(2): 1-15.

[3] Li G and Liu F. Application of a clustering method on sentiment analysis. *Journal of Information Science* 2012; 38(2): 127-139.

[4] Litvin SW, Goldsmith RE and Pan B. Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 2008; 29(3): 458-468.

[5] Bronner F and de Hoog R. Vacationers and eWOM: Who posts, and why, where, and what?. *Journal of Travel Research* 2010; 50(1): 15-26.

[6] Ott M, Choi Y, Cardie C and Hancock JT. Finding deceptive opinion spam by any stretch of the imagination. In: *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. California: AAAI, 2011, pp. 309-319.

[7] Fusilier DH, Montes-y-Gómez M, Rosso P and Cabrera RG. Detecting positive and negative deceptive opinions using PU-learning. *Information Processing & Management* 2015; 51(4): 433-443.

[8] Yoo KH and Gretzel U. Comparison of deceptive and truthful travel reviews. In: W Höpken, U Gretzel and R Law (Eds.), *Information and Communication Technologies in Tourism*. Vienna: Springer, 2009, pp. 37-47.

[9] Jindal N and Liu B. Opinion spam and analysis. In: *Proceedings of the International Conference on Web Search and Data Mining*. New York: ACM, 2008, pp. 219-230.

[10] Johnson MK and Raye CL. Reality monitoring. *Psychological Review* 1981; 88(1): 67-85.

[11] Heydari A, ali Tavakoli M, Salim N and Heydari Z. Detection of review spam: A survey. *Expert Systems with Applications* 2015; 42(7): 3634-3642.

[12] Vrij A, Edward K, Roberts KP and Bull R. Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior* 2000; 24(4): 239-263.

[13] Bond CF and DePaulo BM. Accuracy of deception judgments. *Personality and Social Psychology Review* 2006; 10(3): 214-234.

[14] Banerjee S and Chua AYK. A linguistic framework to distinguish between genuine and deceptive online reviews. In: *Proceedings of the International Conference on Internet Computing and Web Services*. Hong Kong: IAENG, 2014, pp. 501-506.

[15] Ghose A and Ipeirotis PG. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering* 2011; 23(10): 1498-1512.

[16] Tausczik YR and Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 2010; 29(1): 24-54.

[17] Ekiz E, Khoo-Lattimore C and Memarzadeh F. Air the anger: Investigating online complaints on luxury hotels. *Journal of Hospitality and Tourism Technology* 2012; 3(2): 96-106.

[18] Hua W, Chan A and Mao Z. Critical success factors and customer expectation in budget hotel segment—A case study of China. *Journal of Quality Assurance in Hospitality & Tourism* 2009; 10(1): 59-74.

[19] Jiang J, Gretzel U and Law R. Influence of star rating and ownership structure on brand image of Mainland China hotels. *Journal of China Tourism Research* 2014; 10(1): 69-94.

[20] Zhou L, Burgoon JK, Twitchell DP, Qin T and Nunamaker Jr JF. A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems* 2004; 20(4): 139-165.

[21] DePaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K and Cooper H. Cues to deception. *Psychological Bulletin* 2003; 129(1): 74-118.

[22] Harris CG. Detecting deceptive opinion spam using human computation. In: *Proceedings for the Workshops on Artificial Intelligence*. California: AAAI, 2012, pp. 87-93.

[23] Afroz S, Brennan M and Greenstadt R. Detecting hoaxes, frauds, and deception in writing style online. In: *Proceedings of the Symposium on Security and Privacy*. New York: IEEE, 2012, pp. 461-475.
[24] Mihalcea R and Burzo M. Towards multimodal deception detection - step 1: Building a collection of deceptive videos. In: *Proceedings of the International Conference on Multimodal Interaction*. New York: ACM, 2012, pp. 189-192.
[25] Newman ML, Pennebaker JW, Berry DS and Richards JM. Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin* 2003; 29(5): 665-675.
[26] Burgoon JK and Qin T. The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology* 2006; 25(1): 76-96.
[27] Zakaluk BL and Samuels SJ. Readability: Its past, present and future. Newark: International Reading Association, 1988.
[28] Chall JS and Dale E. Readability revisited: The new Dale-Chall readability formula. Cambridge: Brookline Books, 1995.
[29] Cao Q, Duan W and Gan Q. Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems* 2011; 50(2): 511-521.
[30] Hancock JT, Curry LE, Goorha S and Woodworth M. On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes* 2007; 45(1): 1-23.
[31] Rayson P, Wilson A and Leech G. Grammatical word class variation within the British National Corpus sampler. *Language and Computers* 2001; 36(1): 295-306.
[32] Godin S. All marketers are liars: The power of telling authentic stories in a low trust world. New York: Penguin, 2005.
[33] Hancock JT, Curry L, Goorha S and Woodworth M. Automated linguistic analysis of deceptive and truthful synchronous computer-mediated communication. In: *Proceedings of the Hawaii International Conference on System Sciences*. New York: IEEE, 2005, pp. 1-10.
[34] Ravid D. Emergence of linguistic complexity in later language development: Evidence from expository text construction. In: D Ravid and H Shyldkrot (Eds.), *Perspectives on Language and Language Development*. The Netherlands: Kluwer Academic, 2005, pp. 337-355.
[35] Hu N, Bose I, Koh NS and Liu L. Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems* 2012; 52(3): 674-684.
[36] Maurer C and Schaich S. Online customer reviews used as complaint management tool. In: R Law, M Fuchs and F. Ricci (Eds.), *Information and Communication Technologies in Tourism*. Vienna: Springer, 2011, pp. 499-511.
[37] Gunsch MA, Brownlow S, Haynes SE and Mabe Z. Differential linguistic content of various forms of political advertising. *Journal of Broadcasting & Electronic Media* 2000; 44(1): 27-42.
[38] Ekman P and Friesen WV. Nonverbal leakage and clues to deception. *Psychiatry: Interpersonal and Biological Processes* 1969; 32(1): 88-106.
[39] Zuckerman M, DePaulo BM and Rosenthal R. Verbal and nonverbal communication of deception. In: L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14). New York: Academic Press, 1981, pp. 1-59.
[40] Hyland K. Writing without conviction? Hedging in science research articles. *Applied Linguistics* 1996; 17(4): 433-454.
[41] Ascaniis SD and Gretzel U. What's in a travel review title?. In: M Fuchs, F Ricci and L Cantoni (Eds.), *Information and Communication Technologies in Tourism*. Vienna: Springer, 2012, pp. 494-505.
[42] Gräbner D, Zanker M, Fliedl G and Fuchs M. Classification of customer reviews based on sentiment analysis. In: M Fuchs, F Ricci and L Cantoni (Eds.), *Information and Communication Technologies in Tourism*. Vienna: Springer, 2012, pp. 460-470.
[43] Jiang J, Gretzel U and Law R. Do negative experiences always lead to dissatisfaction? - Testing attribution theory in the context of online travel reviews. In: U Gretzel, R Law and M Fuchs (Eds.), *Information and Communication Technologies in Tourism*. Vienna: Springer, 2010, pp. 297-308.
[44] Basiri ME, Ghasem-Aghaee N and Naghsh-Nilchi AR. Exploiting reviewers' comment histories for sentiment analysis. *Journal of Information Science* 2014; 40(3): 313-328.
[45] Gerdes Jr J, Stringam BB and Brookshire RG. An integrative approach to assess qualitative and quantitative consumer feedback. *Electronic Commerce Research* 2008; 8(4): 217-234.
[46] Dawes J. Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. *International Journal of Market Research* (2008; accessed February 2013); 51(1). http://ssrn.com/abstract=2013613
[47] Ghiselli EE. All or none versus graded response questionnaires. *Journal of Applied Psychology* 1939; 23(3): 405-415.
[48] Dawes J. Five point vs eleven point scales: Does it make a difference to data characteristics?. *Australasian Journal of Market Research* 2002; 10(1): 39-47.
[49] Gretzel U, Yoo KH and Purifoy M. Online travel reviews study: Role and impact of online travel reviews. A&M University, Texas: Laboratory for Intelligent Systems in Tourism, 2007.
[50] Anderson J. LIX and RIX: Variations on a little-known readability index. *Journal of Reading* 1983; 26: 490-496.
[51] Bjornsson CH. Readability of newspapers in 11 languages. *Reading Research Quarterly* 1983; 18(4): 480-497.
[52] Korfiatis N, García-Bariocanal E and Sánchez-Alonso S. Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications* 2012; 11(3): 205-217.

[53] Bond GD and Lee AY. Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology* 2005; 19(3): 313-329.
[54] Zhou L and Zhang D. Following linguistic footprints: Automatic deception detection in online communication. *Communications of the ACM* 2008; 51(9): 119-122.
[55] Ott M, Cardie C and Hancock JT. Negative deceptive opinion spam. In: *North American Chapter of the Association for Computational Linguistics*. Pennsylvania: ACL, 2013, pp. 497-501.
[56] Keshtkar F and Inkpen D. A hierarchical approach to mood classification in blogs. *Natural Language Engineering* 2012; 18(1): 61-81.
[57] Tsur O, Davidov D and Rappoport A. A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In: *Proceedings of the International Conference on Weblogs and Social Media*. California: AAAI, 2010, pp. 162-169.
[58] Pasupathi M. Telling and the remembered self: Linguistic differences in memories for previously disclosed and previously undisclosed events. *Memory* 2007; 15(3): 258-270.
[59] Boals A and Klein K. Word use in emotional narratives about failed romantic relationships and subsequent mental health. *Journal of Language and Social Psychology* 2005; 24(3): 252-268.
[60] Lee CC, Welker RB and Odom MD. Features of computer-mediated, text-based messages that support automatable, linguistics-based indicators for deception detection. *Journal of Information Systems* 2009; 23(1): 19-24.
[61] Pennebaker JW, Booth RJ and Francis ME. Linguistic Inquiry and Word Count [Software]. Austin, TX: LIWC.net, 2007.
[62] Klein D and Manning CD. Accurate unlexicalized parsing. In: *Proceedings of the Annual Meeting on Association for Computational Linguistics*. Pennsylvania: ACL, 2003, pp. 423-430.
[63] Menard S. Applied logistic regression analysis: Quantitative applications in the social sciences. Beverley Hills: Sage University Press, 2002.
[64] Pallant J. SPSS survival manual: A step by step guide to data analysis using SPSS for Windows version 10. Buckingham: Open University Press, 2001.
[65] Sarkar TD. Impact of online interactivity dimensions on library website quality. *Annals of Library and Information Studies* 2013; 59(4): 231-239.
[66] Otterbacher J. Gender, writing and ranking in review forums: A case study of the IMDb. *Knowledge and Information Systems* 2013; 35(3): 645-664.
[67] Blooma MJ, Goh DHL and Chua AYK. Predictors of high-quality answers. *Online Information Review* 2012; 36(3): 383-400.
[68] Matzat U. Disciplinary differences in the use of internet discussion groups: Differential communication needs or trust problems?. *Journal of Information Science* 2009; 35(5): 613-631.
[69] Oh O, Agrawal M and Rao HR. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly* 2013; 37(2): 407-426.
[70] Licht MH. Multiple regression and correlation. In: G Laurence and YR Paul (Eds.), *Reading and understanding multivariate statistics*. Washington: APA, 1995, pp. 19-64.
[71] O'Brien RM. A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 2007; 41(5): 673-690.
[72] Forman C, Ghose A and Wiesenfeld B. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research* 2008; 19(3): 291-313.
[73] Wu PF, van der Heijden H and Korfiatis N. The influences of negativity and review quality on the helpfulness of online reviews. In: *Proceedings of the International Conference on Information Systems*. http://epubs.surrey.ac.uk/7533/2/icis2011.final.pdf (2011, accessed June 2015), pp. 1-10.
[74] Buller DB and Burgoon JK. Interpersonal deception theory. *Communication Theory* 1996; 6(3): 203-242.
[75] Hunt KM. Gaming the system: Fake online reviews v. consumer law. *Computer Law & Security Review* 2015; 31(1): 3-25.
[76] Petz G, Karpowicz M, Fürschuß H, Auinger A, Stříteský V and Holzinger A. Computational approaches for mining user's opinions on the Web 2.0. *Information Processing & Management* 2014; 50(6): 899-908.