Study of authentic and fictitious online reviews

Snehasish, Banerjee

2017

Snehasish, B. (2017). Study of authentic and fictitious online reviews. Doctoral thesis,
Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/89391>

<https://doi.org/10.32657/10220/46219>

**STUDY OF
AUTHENTIC AND FICTITIOUS ONLINE REVIEWS**

BANERJEE SNEHASISH

WEE KIM WEE SCHOOL OF COMMUNICATION AND INFORMATION

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2017

ACKNOWLEDGMENTS

I take this opportunity to express my heartfelt gratitude to my supervisor Dr. Alton Y. K. Chua for his efforts and guidance in shaping my research. He has been the perfect antidote to all challenges that I have faced in my PhD life.

I would like to thank the members of my Thesis Advisory Committee, Dr. Jung-Jae Kim, and Dr. Lee Chu Keong. Their comments and suggestions helped me think about my research from different perspectives.

I am grateful to all staff in the Wee Kim Wee School of Communication and Information for their help. In particular, I really appreciate the assistance from Eunice Chua, Jesslyn Tan, Joanne Quek, Juliza Ibrahim, and Kris Ang.

Thanks to all participants and annotators involved in my research. Special thanks to: Anjan Pal, Aravind Raamkumar, Chew Loo Hong, Dennis Linders, Dominic Heng, Fendy Kosnatha, Ivy Sarkar, Jayakumar Santhosh, Kaur Avneet, Khasfariyati Razikin, Lalla Priya Hareesh, Loo Chia Feng, Lucia Bernarda, Mohamed Ariffin, Radhika Shenoy, Paudel Sunil, Rao Sonia, Reza Chowdhury, Sim Kim Heng, Soumodeep Dutta, Sowmyalatha, Su Shwe Zin Thein, Tang Kah Hee, Wang Bo, and Yanru Guo.

Last but not the least, I am eternally grateful to my parents. They have offered unwavering support and encouragement through all my endeavors. Their constant blessing is the key to my success. Sharing my stories of struggle and achievement along the way with them over many hours of phone calls and Skype sessions made my tough PhD journey endurable.

Finally, thank you for reading this section.

ABSTRACT

This research investigates the ways in which differences between authentic and fictitious reviews on the Internet are related to users' perceived review authenticity. It addresses two research questions. The first deals with actual review authenticity while the second focuses on humans' perceived review authenticity. Informed by the literature on deception and information processing, a theoretical model is developed.

Guided by the model, this research examines actual differences between authentic and fictitious reviews using computational techniques. A dataset of 1,800 reviews (900 authentic + 900 fictitious) was used for investigation. Next, it examines humans' perceived differences between authentic and fictitious reviews by surveying 380 participants.

This research finds that authentic reviews are different from fictitious entries in terms of the linguistic cues of exaggeration and specificity. Moreover, humans' perceptions of exaggeration and specificity are related to perceived authenticity. The relations tend to be particularly strong among individuals who believe that the authenticity of information cannot be justified easily. However, humans remain generally sub-par in discerning review authenticity.

Overall, this research shows that even though authentic and fictitious reviews are distinguishable, human ability to discern authenticity is not overly promising. Therefore, it develops a guideline based on the actual differences between authentic and fictitious reviews to teach humans how to discern authenticity. The guideline is used as an intervention in an experimental design involving 240 participants. It is found to improve human ability to perceive differences between authentic and fictitious reviews.

PUBLICATIONS

Content of this thesis appears partially in the following publications:

Journal Papers

Banerjee, S., Chua, A. Y. K., & Kim, J. J. (2017). Don't be deceived: Using linguistic analysis to learn how to discern online review authenticity. *Journal of the Association for Information Science and Technology*, 68(6), 1525-1538.

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.

Banerjee, S., & Chua, A. Y. K. (2017). Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels. *Journal of Information Science*, 43(1), 122-134.

Banerjee, S., & Chua, A. Y. K. (2016). In search of patterns among travellers' hotel ratings in TripAdvisor. *Tourism Management*, 53, 125-131.

Banerjee, S., & Chua, A. Y. K. (2014). Dissecting genuine and deceptive kudos: The case of online hotel reviews. *International Journal of Advanced Computer Science and Applications*, Special Issue on Extended Papers from SAI Conference 2014, 4(3), 28-35.

Banerjee, S., & Chua, A. Y. K. (2014). A theoretical framework to identify authentic online reviews. *Online Information Review*, 38(5), 634-649. [**Highly Commended Paper Award**]

Conference Papers

Banerjee, S., Chua, A. Y. K., & Kim, J. J. (2015). Let's vote to classify authentic and manipulative online reviews: The role of comprehensibility, informativeness and writing style. *Proceedings of the Science and Information Conference* (pp. 77-83). New York, NY: IEEE.

Banerjee, S., Chua, A. Y. K., & Kim, J. J. (2015). Distinguishing between authentic and fictitious user-generated hotel reviews. *Proceedings of the International Conference on Computing Communication and Networking Technologies* (pp. 1-7). New York, NY: IEEE.

Banerjee, S., Chua, A. Y. K., & Kim, J. J. (2015). Using supervised learning to classify authentic and fake online reviews. *Proceedings of the International Conference on Ubiquitous Information Management and Communication* (pp. 88:1-7). New York, NY: ACM.

Banerjee, S., & Chua, A. Y. K. (2014). Applauses in hotel reviews: Genuine or deceptive? *Proceedings of the Science and Information Conference* (pp. 938-942). New York, NY: IEEE.

Banerjee, S., & Chua, A. Y. K. (2014). A linguistic framework to distinguish between genuine and deceptive online reviews. *Proceedings of the International Conference on Internet Computing and Web Services* (pp. 501-506). Hong Kong: IAENG. [**Certificate of Merit Award**]

Banerjee, S., & Chua, A. Y. K. (2014). A study of manipulative and authentic negative reviews. *Proceedings of the International Conference on Ubiquitous Information Management and Communication* (pp. 76:1-6). New York, NY: ACM.

TABLE OF CONTENTS

CHAPTER ONE INTRODUCTION	1
1.1. Research Motivation	1
1.1.1. Problem Statement.....	1
1.1.2. Gaps in Current Works	3
1.2. Research Goal	6
1.3. Methodological Outline	9
1.4. Potential Contributions	11
1.4.1. Potential Theoretical Contributions.....	11
1.4.2. Potential Practical Contributions	14
1.5. Research Delimitations	14
1.6. Organization of the Thesis	16
CHAPTER TWO LITERATURE REVIEW	18
2.1. Authentic versus Fictitious Reviews	19
2.2. Users' Perceptions of Reviews	21
2.3. Theoretical Model.....	23
2.3.1. Framework of Linguistic Cues	25
2.3.2. Framework of User Perceptions	31
2.4. Summary of the Literature Review.....	36
CHAPTER THREE LINGUISTIC STUDY – METHODS.....	37
3.1. Ground Truth	38
3.2. Data Sources	41
3.2.1. Review Websites	42
3.2.2. Hotels.....	43
3.3. Data Collection	45
3.3.1. Authentic Reviews.....	46
3.3.2. Fictitious Reviews	50
3.4. Operationalization and Measurement	59
3.5. Data Analysis.....	67
3.6. Summary of the Linguistic Study – Methods	73

CHAPTER FOUR	LINGUISTIC STUDY – FINDINGS.....	75
4.1.	Descriptive Statistics of the Linguistic Dataset	76
4.2.	Distinguishing Authentic from Fictitious Reviews	80
4.3.	Unique Linguistic Traits	81
4.3.1.	Comprehensibility	83
4.3.2.	Exaggeration.....	85
4.3.3.	Specificity	90
4.3.4.	Tentativeness	93
4.4.	Comparing Linguistic Study Results with Prior Works	96
4.5.	Summary of the Linguistic Study – Findings	103
CHAPTER FIVE	USER STUDY – METHODS.....	105
5.1.	Survey Stimuli	107
5.1.1.	Simulated Review Website.....	108
5.1.2.	Choice of Reviews.....	110
5.2.	Survey Procedure	112
5.3.	Questionnaire Development	117
5.3.1.	Measures	118
5.3.2.	Evaluation of the Measures	123
5.4.	Analytical Approach	126
5.4.1.	Quantitative Analysis	126
5.4.2.	Qualitative Analysis	131
5.5.	Methods of the Robustness Check.....	134
5.6.	Summary of the User Study – Methods.....	140
CHAPTER SIX	USER STUDY – FINDINGS.....	142
6.1.	Descriptive Statistics of the User Dataset.....	144
6.2.	Role of Perceived Linguistic Cues	151
6.3.	Moderating Role of Epistemic Beliefs	154
6.3.1.	Perceived Reliability of Knowledge.....	154
6.3.2.	Perceived Justification for Knowing	157
6.4.	Results of the Qualitative Analysis.....	161

6.5. Results of the Robustness Check	170
6.6. Comparing User Study Results with Prior Works	172
6.7. Summary of the User Study – Findings	176
CHAPTER SEVEN CONCLUSION.....	177
7.1. Summary of the Research Findings	179
7.2. Contributions	182
7.2.1. Theoretical Contributions	182
7.2.2. Practical Contributions	189
7.3. Ethical Considerations	193
7.4. Limitations	194
7.5. Future Research Directions.....	195
Appendix A Identified Hotels	198
Appendix B Instructions for Fictitious Reviews	200
Appendix C Incorrectly Annotated Reviews.....	202
Appendix D Detailed Classification Results	203
Appendix E Screenshots of the Simulated Website	204
Appendix F Advertisement for Study Participation.....	206
Appendix G Guideline for the Robustness Check	207
Appendix H Details of the User Dataset	209
Appendix I Participants’ Quotations	213
References	214

LIST OF FIGURES

1.1. Research overview	9
1.2. Organization of the chapters	17
2.1. Overview of Chapter Two	19
2.2. The theoretical model of authentic and fictitious reviews (TAF)	35
3.1. Overview of Chapter Three	38
3.2. Overview of the data collection procedure	46
3.3. Process of collecting authentic reviews for a given hotel	49
3.4. Process of collecting fictitious reviews for a given hotel	56
3.5. Iterations of the five-fold cross-validation	68
3.6. Summary of the analytical procedure for the Linguistic Study	73
4.1. Overview of Chapter Four	76
5.1. Overview of Chapter Five	107
5.2. Design template used for the simulated review website	110
5.3. Three-step process to choose reviews as the stimuli	112
5.4. Summary of the analytical procedure for the User Study	133
5.5. Decision-tree format of the guideline to discern review authenticity	136
5.6. Tasks in the Robustness Check	140
6.1. Overview of Chapter Six	144
6.2. PLS result showing the role of perceived linguistic cues	152
6.3. Multi-group PLS result for perceived reliability of knowledge	157
6.4. Multi-group PLS result for perceived justification for knowing	159
7.1. Overview of Chapter Seven	179
7.2. Diagrammatic representation of the results of the TAF	184

LIST OF TABLES

2.1. Works on authentic versus fictitious reviews	21
2.2. Works on users' perceptions of reviews	23
2.3. Linguistic cues and linguistic traits.....	31
2.4. Summary of the TAF	36
3.1. Scheme to assign review polarities	48
3.2. Agreement between ratings and annotations for authentic reviews.....	50
3.3. Agreement between ratings and annotations for fictitious reviews	57
3.4. Distribution of reviews across hotel categories and review polarities.....	58
3.5. Operationalization of the linguistic cues.....	64
3.6. Pruned set of variables	66
3.7. Confusion matrix of classification possibilities	68
4.1. Descriptive statistics with respect to comprehensibility	77
4.2. Descriptive statistics with respect to exaggeration	78
4.3. Descriptive statistics with respect to specificity	79
4.4. Descriptive statistics with respect to tentativeness	80
4.5. Comparing the framework of linguistic cues with the baselines	81
4.6. Information gain and Chi-squared feature selection	82
4.7. Linguistic differences based on comprehensibility.....	85
4.8. Linguistic differences based on exaggeration.....	90
4.9. Linguistic differences based on specificity	93
4.10. Linguistic differences based on tentativeness	95
4.11. Summary of the results pertaining to the framework of linguistic cues	96
4.12. Linguistic differences between authentic and fictitious reviews	103
5.1. Reviews used as stimuli across hotel categories and review polarities	108
5.2. Distribution of reviews across versions of LoveToTravel.com.....	114
5.3. Distribution of participants across versions of LoveToTravel.com.....	117
5.4. Items in the first segment of the questionnaire	121
5.5. Items in the second segment of the questionnaire	123
5.6. Methods employed for evaluation of measures	126

5.7. Differences consistently prominent across review titles and descriptions.....	135
6.1. Demographic details of the participants	145
6.2. Descriptive statistics of items in the first segment of the questionnaire	146
6.3. Descriptive statistics of items in the second segment of the questionnaire	147
6.4. Evaluation of formative constructs	149
6.5. Evaluation of reflective constructs.....	149
6.6. Results of evaluating the measures	150
6.7. Moderating role played by perceived justification for knowing.....	160
6.8. Coding scheme for the qualitative analysis	162
6.9. Summary of the results pertaining to the framework of user perceptions	169
7.1. Summary of the results pertaining to the TAF.....	183
7.2. Summary of the research contributions	193

CHAPTER ONE INTRODUCTION

Social media applications are charting a meteoric rise in popularity. They are used daily by about 70% of Internet users (Pew Research Centre, 2015), who spawn enormous volumes of user-generated content such as edits on Wikipedia, status updates on Facebook, and tweets on Twitter.

Another form of user-generated content that has mushroomed rapidly includes online reviews for products or services. Posted on review websites such as TripAdvisor.com, reviews enable users to share their post-purchase experiences with other potential buyers in the online community.

Users often rely on reviews to decide what to buy, and what to avoid. They consider reviews as having been posted free from commercial interests (Gesenhues, 2013). Given the supposed proximity of reviews with sentiment on the ground, they assume reviews to be more authentic compared with marketer-generated information in the mainstream media (Filieri, 2016; Gu et al., 2012; Miao et al., 2010). Unsurprisingly, about 95% of buyers read reviews prior to making purchases while nearly 89% consider reviews as being influential in their decision-making (Hu & Chen, 2016).

1.1. Research Motivation

1.1.1. Problem Statement

The authenticity of reviews cannot be taken for granted due to what is known as spam 2.0. In general, spam refers to the abuse of electronic messaging systems through

dissemination of unsolicited messages in bulk (Whitworth & Whitworth, 2004). Spam 2.0 specifically refers to the propagation of unsolicited contents to infiltrate social media applications (Hayati et al., 2010). Since such unsolicited contents are deliberately written to be passed off as authentic, they are difficult to be detected by state-of-the-art countermeasures such as blacklisting and keyword filtering (Blanzieri & Bryl, 2008; Ridzuan et al., 2010).

In the context of review websites, spam 2.0 includes posting fictitious reviews by mimicking authentic ones. Review websites are unfortunately peppered with fictitious reviews. Hotel managers have been caught asking employees to post fictitious positive reviews to maliciously boost the ratings of their properties (Allen, 2012). Businesses often offer discounts in exchange of fictitious positive reviews, which could be posted by users without necessarily harboring malicious intentions (Anderson & Simester, 2014). Ironically, extolling a product with fictitious positive reviews continuously for about 45 days on the review website Amazon.com elevates it to the numero uno bestseller rank in its category (The Guardian, 2011).

To aggravate the problem, the polarity of fictitious reviews is not always positive. Reviews of negative and moderate polarities can also be fictitious. Posting negative fictitious reviews is known to be encouraged by businesses as a way to slander competitors (Kerr, 2013). Fictitious reviews may also express a moderate tone, neither too glowing nor too critical, to appear realistic (Allen, 2012). Perhaps unsurprisingly, as much as 30% of all reviews—positive, negative and moderate alike—posted on the review website Yelp.com are known to be fictitious (Vega, 2012).

The undeniable existence of foul play suggests that reviews should be taken with a pinch of salt. While some reviews are authentic accounts of post-purchase experiences, others are fictitious tales based on imagination. Moreover, since fictitious reviews are written to resemble authentic ones, differences between the two are difficult to detect. This leaves users clutching at straws to discern review authenticity. When they fail to do so, their perceptions of products and services stand a chance to be insidiously distorted. Hence, they run the risk of being deceived. The risk is especially high given users' default tendency to consider reviews as being authentic (Sidali et al., 2009; Vrij & Baxter, 1999).

Whether reviews are taken as gospel truth or viewed with suspicion is crucial for both users and businesses. When users perceive fictitious reviews as being authentic, they might end up purchasing inferior products and services while overlooking superior ones. Conversely, users being suspicious of reviews that are authentic defeat the whole purpose of the review system. Furthermore, users' distorted perceptions of products and services could make or break the fate of businesses. While the virtual image of some businesses is unfairly sullied, that of others might be illegitimately enhanced. Therefore, the problem posed by the prevalence of fictitious reviews amid authentic ones on the Internet warrants urgent scholarly attention.

1.1.2. Gaps in Current Works

In response to the problem, scholarly efforts continue to spawn a growing body of works that can be grouped into two clusters. One cluster includes works that distinguish between authentic and fictitious reviews. The other cluster includes works that examine users' perceptions of reviews.

A research gap is identified in each of the two clusters. Works in the first cluster that distinguish between authentic and fictitious reviews lack an overarching theoretical framework. Some works relied on review-centric, reviewer-centric and product-centric metadata features captured automatically on websites (Jindal & Liu, 2008). Others leveraged on the possible part-of-speech tags, unigrams, bigrams, trigrams, and several psycholinguistic dimensions (Ott et al., 2011). However, scholars appear unsure about the width of the net to cast in order to identify differences between authentic and fictitious reviews. For example, while some narrowly relied on linguistic cues such as exaggeration (Yoo & Gretzel, 2009), others at the extreme end used almost a brute-force strategy by including a plethora of variables, thereby compromising parsimony (Ott et al., 2011). Therefore, as suggested by Hauch et al. (2015), this scholarly terrain seems to deserve an investigation that will develop a comprehensive yet parsimonious list of variables to distinguish between authentic and fictitious reviews.

This research considers that the textual nature of reviews could offer a vantage perspective to ascertain review authenticity. This is because empirical and anecdotal evidences have long suggested that the cognitive states differ between individuals writing authentic texts, and those writing fictitious texts. The differences in the cognitive states reflect in individuals' language behavior. Hence, texts describing real experiences differ from those concocted out of imagination (Johnson & Raye, 1981; McQuaid et al., 2015; Porter & ten Brinke, 2010; Undeutsch, 1967; Vrij et al., 2000). Unsurprisingly, review texts continue to be heralded as “the first thing to be considered” to predict authenticity (Heydari et al., 2015, p. 3635). Yet, there does not exist a comprehensive set of textual variables that could help distinguish between authentic and fictitious reviews.

Another research gap is identified among works in the second cluster that examine users' perceptions of reviews. These works shed little light on the extent to which users are able to distinguish between authentic and fictitious reviews. The focus has often been on users' perceptions of review authenticity albeit implicitly. For example, several works examined how reviews influenced perceived credibility (Qiu et al., 2012), perceived usefulness (Casaló et al., 2011), and intention to trust (Sidali et al., 2009). However, reviews were implicitly assumed to be authentic. Hence, the extent to which users are able to identify textual differences between authentic and fictitious reviews in order to discern review authenticity remains unknown (Toma & D'Angelo, 2015). This calls for investigating the ways in which users' ability to identify textual differences between authentic and fictitious entries help them discern review authenticity.

In this vein, there is evidence that users perceive most reviews as being authentic due to their inherent truth bias (Gu et al., 2012; Ott et al., 2011; Vrij & Baxter, 1999; Vrij, 2008). Nonetheless, the extent to which they are susceptible to truth bias could stem from their epistemic beliefs that include perceptions of the reliability of knowledge, and the justification for knowing (Hofer, 2004; Hofer & Pintrich, 1997). Some individuals perceive knowledge as being relatively reliable, and justify the process of knowing based on gut feeling. In contrast, others take knowledge with a pinch of salt, and justify the process of knowing through rigorous rules of inquiry (Kammerer et al., 2013; Qiu et al., 2012; Tu et al., 2008). Yet, little is known about the extent to which users with different epistemic beliefs are able to perceive textual differences between authentic and fictitious entries in order to discern review authenticity.

1.2. Research Goal

Informed by the problem statement and the research gaps, this research seeks to investigate the ways in which *linguistic differences* between authentic and fictitious reviews are related to users' *perceived review authenticity*. For this purpose, it draws from the theoretical perspectives of deception and information processing. The former informs the study of linguistic differences between authentic and fictitious reviews while the latter guides the study of users' perceptions. Inspired by these theoretical perspectives, this research develops a *theoretical model of authentic and fictitious reviews (TAF)*. Moving down the ladder of abstraction, the model is conceived as a *framework of linguistic cues* to distinguish between authentic and fictitious reviews, and a *framework of user perceptions* for authentic and fictitious reviews.

This research uses reviews for hotels as the test case for investigation. *Authentic reviews* are defined as those contributed by users who had paid, patronized and experienced a given hotel. On the other hand, *fictitious reviews* refer to those written by participants in a research setting based on imagination without any post-purchase experience of staying in the hotel.

Two research questions (RQs) are formulated to accomplish the research goal. The first RQ investigates actual linguistic differences between authentic and fictitious reviews. The second RQ investigates users' perceived authenticity of reviews. Each RQ is expressed as two subordinate research objectives (ROs), the second of which operates at a deeper level of investigation than the first. The RQs and the ROs are as follows.

RQ 1: *How do authentic reviews linguistically differ from fictitious entries?*

RO 1.1: To distinguish between authentic and fictitious reviews based on linguistic cues.

RO 1.2: To identify specific linguistic traits unique to authentic and fictitious reviews.

Investigation of this RQ is informed by the *framework of linguistic cues* to distinguish between authentic and fictitious reviews. *Linguistic cues* refer to the nature of the textual content used in reviews (Newman et al., 2003) while *linguistic traits* refer to granular properties of reviews in terms of a given linguistic cue (Pérez-Sabater et al., 2008). An example of a linguistic cue, which is made up of multiple linguistic traits, is exaggeration. Emotion words, and hyperbolic expressions are examples of linguistic traits corresponding to exaggeration.

Addressing RO 1.1 is significant because prior works advocate the use of linguistic cues to distinguish between authentic and fictitious reviews (Harris, 2012; Vartapetian & Gillam, 2012; Yoo & Gretzel, 2009). However, the literature currently lacks an overarching framework for this purpose. By leaning on previous works, this research identifies a possible set of linguistic cues along with linguistic traits that is comprehensive yet parsimonious to ascertain review authenticity.

Furthermore, addressing RO 1.2, which delves deeper into the results of RO 1.1, is significant because prior works often emphasized on accurate classification of authentic and fictitious reviews without necessarily engaging in finer-grained analyses of their specific linguistic traits (O'Mahony & Smyth, 2009; Ott et al., 2011; Zhou et al., 2008). As a result, nuances in linguistic traits as a function of review authenticity remain largely unclear. This murkiness in the literature surrounding possible linguistic differences between authentic and fictitious reviews is addressed in this research.

RQ 2: *To what extent are users able to perceive the authenticity of reviews?*

RO 2.1: To analyze how perceived linguistic cues are related to perceived authenticity.

RO 2.2: To examine how epistemic beliefs affect the relation between perceived linguistic cues and perceived authenticity.

Investigation of this RQ is informed by the *framework of user perceptions* for authentic and fictitious reviews. *Perceived authenticity* refers to the extent to which users consider a review to have been written after real experiences (Laing et al., 2014).

Perceived linguistic cues refer to the extent to which users are able to identify textual nuances that help distinguish between authentic and fictitious reviews in reality (Toma & D'Angelo, 2015). *Epistemic beliefs* refer to users' perceptions of the reliability of knowledge, and the justification for knowing (Hofer & Pintrich, 1997).

Addressing RO 2.1 is significant because prior works have shed little light on the extent to which users' perceived linguistic cues of reviews inform their ability to discern review authenticity. Nonetheless, users are known to perceive most reviews as being authentic given their inherent truth bias (Gu et al., 2012; Vrij & Baxter, 1999; Vrij, 2008). Therefore, this research explores the extent to which actual linguistic differences between authentic and fictitious reviews are perceived by users to discern authenticity.

Furthermore, addressing RO 2.2, which delves deeper into the results of RO 2.1, is significant because the literature is currently mum on the extent to which users' perceptions of authenticity could be affected by their epistemic beliefs. In particular, epistemologically naïve users could differ in their perceptions from epistemologically robust individuals (Kammerer et al., 2013; Qiu et al., 2012; Tu et al., 2008). Hence, this

research sheds light on how epistemic beliefs moderate the relation between perceived linguistic cues and perceived authenticity. Figure 1.1 offers an overview of this research.

Figure 1.1: Research overview.

1.3. Methodological Outline

This research conducted two related studies followed by a Robustness Check. The first study addressed RQ 1 that deals with actual linguistic differences between authentic

and fictitious reviews. It is referred as the Linguistic Study. The second study addressed RQ 2 that focuses on users' perceived authenticity of reviews. It is referred as the User Study. To confirm the findings of this research, the User Study was followed by an additional examination referred as the Robustness Check.

As indicated earlier in Section 1.2 (p. 6), reviews for hotels were used for investigation. This was due to two reasons. First, prior works have shown that hotel reviews constitute a viable corpus to investigate actual as well as perceived differences between authentic and fictitious reviews (Ott et al., 2013; Sidali et al., 2009). Second, the hotel industry represents one of world's fastest growing economic sectors (Fotis et al., 2011). Hotel reviews have also been soaring in popularity (Fotis et al., 2011; Kwon & Sung, 2012). Hence, analyzing authentic and fictitious reviews for hotels not only dovetails the literature but might also offer important implications for practice.

To enhance representativeness and ensure a well-balanced dataset, this research uses datasets that cross three categories of hotels—luxury, budget and mid-range—with three polarities of reviews—positive, negative and moderate. Hotel categories were determined based on hotels' website-assigned ratings. Review polarities were determined based on reviews' ratings. More details on identifying hotel categories and review polarities are described later in Section 3.2.1 (p. 43) and Section 3.3.1 (p. 47) respectively.

For the Linguistic Study, a dataset of authentic and fictitious reviews was created. Authentic reviews were drawn from authenticated websites that allow submission of entries only after valid purchases and stays. Fictitious reviews were solicited from participants. The possible linguistic cues that could help ascertain review authenticity were operationalized. Classification algorithms were employed to identify. Statistical

analyses were conducted to identify the linguistic traits of authentic and fictitious reviews. The quantitative findings were complemented with qualitative insights from the dataset.

The Linguistic Study was followed by the User Study, which recruited participants for an online survey. Participants' perceived linguistic cues were examined for a set of reviews—some authentic, and others fictitious. The survey questionnaire sought closed-ended and open-ended responses. Quantitative analysis involved structural equation modeling while qualitative analysis included content analysis. The extent to which users' perceived linguistic cues were related to perceived authenticity of reviews was examined. The potential moderating role of participants' epistemic beliefs was also investigated.

To further confirm the findings arising from the two studies, this research conducted the Robustness Check. The purpose was to verify if the actual differences between authentic and fictitious reviews—identified in the Linguistic Study—could be taught to humans so as to improve their ability to discern review authenticity compared with their default ability—as identified in the User Study. Hence, the results of the Linguistic Study were fed as input to develop a guideline on how to identify fictitious reviews amid authentic ones. The guideline was used as an intervention in an experimental design to analyze its efficacy.

1.4. Potential Contributions

1.4.1. Potential Theoretical Contributions

This research has the potential to contribute to the literature in at least three ways. First, it represents one of the earliest attempts to integrate two disparate strands of works on reviews. So far, scholars analyzing differences between authentic and fictitious reviews

have largely ignored users' perceptions (Abulaish & Bhat, 2015; Jindal & Liu, 2008; Ott et al., 2011). Conversely, scholars investigating users' perceptions of reviews have largely neglected whether users are able to discern review authenticity in the first place (Casaló et al., 2011; Qiu et al., 2012; Sidali et al., 2009). Few works hitherto have attempted to bridge the chasm between the two research strands.

This research makes an effort to address the piecemeal scholarship by incorporating the merits of both the strands to address its goal. Specifically, drawing from the first strand, it recognizes the importance of distinguishing between authentic and fictitious reviews. Moreover, informed by the second strand, it appreciates the importance of examining users' perceptions. Inspired by such a twin-perspective, this research ventures into an uncharted territory by investigating the ways in which the actual differences between authentic and fictitious reviews are perceived by users.

Second, this research seeks to develop a framework of linguistic cues to distinguish between authentic and fictitious reviews. Although authentic and fictitious reviews are difficult to detect, some telltale linguistic cues might help in their identification (Johnson & Raye, 1981; McQuaid et al., 2015). Siding with this argument, some works have investigated how authentic and fictitious reviews could be distinguished linguistically (Harris, 2012; Ott et al., 2011; 2013). However, the literature still lacks an overarching framework that is comprehensive yet parsimonious.

To cover this gap in the literature, this research draws from previous works to develop a framework of linguistic cues in order to ascertain review authenticity (cf. Section 2.3, p. 40). Unlike works that mostly focus on accurate classification of authentic and fictitious reviews without a fine-grained analysis (O'Mahony & Smyth, 2009; Ott et

al., 2011; Zhou et al., 2008), this research aims to use the classification results as a springboard to identify specific linguistic traits of the two. Thus, it can dovetail the current understanding of linguistic differences between authentic and fictitious reviews.

Third, this research seeks to develop a framework of user perceptions for authentic and fictitious reviews. Prior works have shown that users are generally inclined to perceive most reviews as being authentic (Gu et al., 2012; Vrij, 2008). These works suggest that users' perceived authenticity of reviews could be generally high.

Concurrently, there are ample evidences hinting that linguistic cues offer telltale signs to distinguish between authentic and fictitious reviews (Johnson & Raye, 1981; McQuaid et al., 2015; Undeutsch, 1967; Vrij et al., 2000). This raises the question whether users are at all able to perceive the linguistic cues that help ascertain review authenticity. However, the literature is currently silent on whether the actual linguistic differences between authentic and fictitious reviews are consistent with users' perceived differences.

Moreover, few related works have shed light on the individual difference of epistemic beliefs, which affect users' ability to evaluate information. Individuals who are epistemologically naïve could be more easily deceived by fictitious information compared with those who are epistemologically robust (Kammerer et al., 2013; Qiu et al., 2012; Tu, Shih & Tsai, 2008). However, such a possibility is yet to be empirically confirmed in the case of authentic and fictitious reviews. Therefore, the framework of user perceptions developed in this research taking into account the role of perceived linguistic cues, perceived authenticity, and epistemic beliefs could push the frontiers of knowledge on users' decision-making based on reviews.

1.4.2. Potential Practical Contributions

This research can offer implications for users, businesses and review websites. For users, it calls for caution in relying on reviews. In addition, its findings have the potential to empower users with some heuristics that could be followed to distinguish between authentic and fictitious reviews with at least moderate levels of accuracy. The findings could also demonstrate the importance of cultivating robust epistemic beliefs for individuals in order to be discerning while processing online information.

For businesses, this research could help uncover the existence of foul play. In the age of social media, businesses often try to improve their performance by keeping a pulse on users' compliments and complaints expressed in reviews. Such efforts would be severely jeopardized if reviews are not authentic in the first place. Therefore, the findings could give options to businesses on how to sieve out noise from reviews.

For review websites, this research offers some options that could be adopted to minimize the impact of fictitious reviews on users' purchase decisions. For example, they could consider taking the help of software companies to explore the possibility of developing automated systems that can identify authentic reviews, and flag off fictitious ones in real-time with reasonable accuracy. Moreover, review websites could include some guidelines in the frequently-asked-questions section as a way to prevent users from letting down their guards while reading reviews.

1.5. Research Delimitations

This research is demarcated by three delimitations. First, its scope is trained on authentic and fictitious reviews in English language for hotels. English is chosen because

it is the most popular language used on the Internet, supported by some 536.60 million (26.8%) users (Internet World Stats, 2013). Moreover, hotel reviews are investigated due to reasons highlighted earlier in Section 1.3 (p. 10). Therefore, analyzing non-English reviews or those for other products and services is beyond the scope of this research.

Second, this research specifically focuses on language to distinguish between authentic and fictitious reviews. This is because language has been shown to help identify deception in a variety of contexts such as businesses' financial disclosures (Humphreys et al., 2011), crime settings (Fuller et al., 2011, 2015), interviews (Burgoon & Qin, 2006), e-commerce websites (Grazioli, 2004; Grazioli & Jarvenpaa, 2003; Xiao & Benbasat, 2011), and online communication (Vishwanath et al., 2011; Zhou, Burgoon, Nunamaker et al., 2004). Moreover, given that review texts are readily available, leveraging language is a viable approach. In contrast, the viability of approaches that rely on other dimensions such as reviewer-centric and product-centric features would be confined to reviews drawn from only specific websites that offer metadata liberally. Therefore, ascertaining review authenticity using non-linguistic dimensions is beyond the scope of this research.

Third, this research collectively treats disinformative as well as misinformative reviews as fictitious entries. Even though the former is written with malicious intentions and the latter for benign reasons, both share the same goal of trying to appear authentic when they are not. The lines are conceivably blurred between whether a contributor posted an entry with malicious intentions or otherwise. Given that disinformative and misinformative reviews tend to be indistinguishable, exhuming differences between the two is beyond the scope of this research.

1.6. Organization of the Thesis

This opening chapter has provided an overview of the problems posed by the prevalence of fictitious reviews amid authentic ones on the Internet. It has also identified two research gaps. Thereafter, it highlighted the research goal, the methodological outline, and the potential contributions. Finally, the chapter presented the delimitations of the research. As shown in Figure 1.2, the remainder of the thesis proceeds as follows.

Chapter Two (pp. 18-36) offers a review of the literature. Specifically, it discusses works on authentic versus fictitious reviews as well as those on users' perceptions. This eventually culminates in the theoretical model of authentic and fictitious reviews.

Chapter Three (pp. 37-74) elaborates the methods for the Linguistic Study. Specifically, it highlights the challenges involved in collecting data for research on authentic and fictitious reviews. Next, the choice of data sources and the methods for data collection are explained. These are followed by a description of the methods for operationalization, measurement and analysis.

Chapter Four (pp. 75-104) presents the findings of the Linguistic Study. Specifically, it summarizes the descriptive statistics, and presents the results pertaining to the research objectives RO 1.1 as well as RO 1.2. Thereafter, the results are discussed in relation to the literature.

Chapter Five (pp. 105-141) elaborates the methods for the User Study. Specifically, it describes the stimuli, the procedure, and the questionnaire for the survey. These are followed by a description of the analytical approaches. The methods employed for the Robustness Check are also described.

Chapter Six (pp. 142-176) presents the findings of the User Study. Specifically, it summarizes the descriptive statistics, and presents the results pertaining to the research objectives RO 2.1 as well as RO 2.2. The results of the Robustness Check are also presented. Thereafter, the results are discussed in relation to the literature.

Chapter Seven (pp. 177-197) brings the thesis to a close. Specifically, it summarizes the findings of the research, and highlights the contributions. The ethical considerations are also indicated. This is followed by a description of the research limitations. Directions for future research are offered.

Figure 1.2: Organization of the chapters.

CHAPTER TWO LITERATURE REVIEW

This chapter offers a review of the literature to develop the *theoretical model of authentic and fictitious reviews (TAF)*. A theoretical model is defined as a group of inter-related perspectives about a research issue at stake (Heller & Campbell, 1981). The TAF is referred to as a theoretical model because it draws from the theoretical perspectives of deception and information processing that inform not only how authentic reviews could differ from fictitious ones but also the ways in which users might process such entries.

Moving down the ladder of abstraction, the TAF is conceived as the *framework of linguistic cues* to distinguish between authentic and fictitious reviews, and the *framework of user perceptions* for authentic and fictitious reviews. A framework is a theoretical tool that helps predict and explain a central construct (Al-Nasser et al., 2013; McColl & Pranger, 1994). The central constructs of the framework of linguistic cues, and the framework of user perceptions are authenticity and perceived authenticity respectively.

As shown in Figure 2.1, the remainder of the chapter is organized into four sections. The first section (p. 19) reviews the literature on differences between authentic and fictitious reviews. The second section (p. 21) reviews prior works that focus on users' perceptions of reviews. The third section (p. 23) explains the construction of the TAF from previous works. Specifically, it presents the framework of linguistic cues to distinguish between authentic and fictitious reviews (p. 25), and the framework of user perceptions for authentic and fictitious reviews (p. 31). Finally, the fourth section (p. 36) summarizes the chapter.

Figure 2.1: Overview of Chapter Two.

2.1. Authentic versus Fictitious Reviews

Grounded on deception theories such as the information manipulation theory (McCornack, 1992) and the reality monitoring theory (Johnson & Raye, 1981), works that distinguish between authentic and fictitious reviews have a technical orientation. They are mostly conducted by scholars from the computer science discipline using algorithmic techniques (Feng et al., 2012; Jindal & Liu, 2008; Massung et al., 2013). They tend to rely on either the bibliometric approach or the engineering approach of investigation. The former involves examining the statistical properties of information while the latter requires developing and testing new ways to process information (Bates, 2005; Brookes, 1968; Small, 1999). In the context of this research, information encompasses website metadata, patterns or language of reviews.

Some of these works relied on website metadata to distinguish between authentic and fictitious reviews. For example, works such as Li et al. (2011) used review-related

metadata such as ratings along with a set of reviewer-related metadata that included aspects such as number of reviews written by reviewers, their popularity rank, their authority scores calculated using the PageRank algorithm, and their deviation scores computed in terms of entropy as well as variances. Besides, works such as Jindal and Liu (2008) considered product-related metadata that included products' price and popularity.

Some other works relied on patterns to distinguish between authentic and fictitious reviews. For example, Wu et al. (2010) suggested that patterns of positive reviews posted soon after a product has attracted a series of negative entries indicate foul play. Feng et al. (2012) conceived fictitious reviews as those that distort the natural distribution of reviews' ratings. Such reviews are more likely to be written by one-time reviewers than regular reviewers. Moreover, the former maintains more extreme tones. Works such as Lim et al. (2010) suggested that when several reviews are posted in a short span of time by a group of reviewers, such a bursty pattern is a telltale sign of manipulation. After all, spam 2.0 campaigns of posting fictitious reviews generally involve coordinated actions taken by several spammers in short spans of time (Abulaish & Bhat, 2015).

Additionally, some other works relied on linguistic cues to distinguish between authentic and fictitious reviews. For example, works such as Yoo and Gretzel (2009) examined the level of exaggeration in authentic and fictitious reviews. Some of these works specifically emphasized on the accurate classification of authentic and fictitious reviews without necessarily identifying the ways in which the two were different. Works such as Ott et al. (2011, 2013) classified reviews using all the possible parts of speech, unigrams, bigrams, trigrams, and psycholinguistic dimensions that included level of details, cues of exaggeration as well as tentativeness while those such as Fusilier et al.

(2015) proposed a positive-unlabeled-learning approach. The themes in the works on distinguishing between authentic and fictitious reviews are summarized in Table 2.1.

Table 2.1: Works on authentic versus fictitious reviews.

Themes	References
Differences between authentic and fictitious reviews in terms of website metadata	Jindal & Liu (2008) Li et al. (2011)
Differences between authentic and fictitious reviews in terms of patterns	Abulaish & Bhat (2015) Wu et al. (2010)
Differences between authentic and fictitious reviews in terms of linguistic cues	Ott et al. (2013) Yoo & Gretzel (2009)

2.2. Users' Perceptions of Reviews

Grounded on information processing theories such as the dual process theory (Deutsch & Gerard, 1955) and the truth bias theory (Vrij & Baxter, 1999), works that examine users' perceptions of reviews have a socio-psychological orientation. They are mostly conducted by social scientists using studies that involve human participants (Casaló et al., 2011; Gupta & Harris, 2010; Tsang & Prendergast, 2009). They tend to rely on the cognitive approach of investigation, which involves analyzing the thought process of individuals as they operate in the world (Bates, 1979, 2005; Belkin, 1990; Ellis, 1989). In the context of this research, operating in the world encompasses the ways in which users read and process reviews.

Some of these works identified factors that predict users' perceptions of reviews. For example, works such as Lim and Van Der Heide (2015) suggested that users' perceived credibility of reviews could be predicted by the profile of reviewers—often referred as source credibility (Savolainen, 2011), and users' familiarity with the platform. Works such as Xu (2014) examined perceived credibility of reviews as a function of the

interplay among reviewers' reputation and profile pictures as well as reviews' ratings. Reviewers' reputation coupled with the interaction between reviewers' profile pictures and reviews' ratings influenced perceived credibility. Several works have documented significant relations between reviews' ratings and users' perceptions of reviews (Chiou et al., 2013; Qiu et al., 2012; Tsang & Prendergast, 2009).

Some other works relied on users' perceptions of reviews to predict behavioral outcomes. For example, McKnight et al. (2002) suggested that users' perceived trustworthiness could predict their intention to follow the advice expressed in reviews. Several works found users' perceived credibility of reviews to be related to their attitude (Casaló et al., 2011; Lim & Van Der Heide, 2015; Sen & Lerman, 2007). Works such as Bambauer-Sachse and Mangold (2013) found that users' perceptions of reviews' ratings could influence their attitude toward the brand under evaluation, specifically leading to brand equity dilution. Casaló et al. (2011) found that perceived credibility and usefulness of reviews predicted users' attitude, which in turn predicted their intention to follow reviews. Several works have documented users' perceived credibility, perceived trustworthiness, and perceived usefulness of reviews to predict their purchase intentions (Chang & Chen, 2008; Park et al., 2007; Pavlou, 2003; Xia & Bechwati, 2008).

A growing body of works focuses on humans' ability to detect deception and discern authenticity. Although deception is widespread in everyday life, individuals are generally sub-par in discerning what is authentic and what is fictitious. Their success could at best be marginally better than chance (Bond & DePaulo, 2006; Toma & Hancock, 2012). Nonetheless, works such as George et al. (2014) offer hope that exercising caution and utilizing available cues could help humans discern information authenticity.

Additionally, recent works on information processing have started to focus on the individual trait of epistemic beliefs, which include perceptions of the reliability of knowledge, and the justification for knowing (Kammerer et al., 2013; Mason & Ariasi, 2010; Whitmire, 2004). This is because reading and processing information involves subjective decision-making based on aspects of individuals' general disposition about knowledge and knowing (Hofer, 2004; Lakey et al., 2008; Lehman et al., 2014). Works such as Tu et al. (2008) found that individuals' epistemic beliefs affect the ways in which they evaluate the credibility of online information. Using eye-tracking, Mason and Ariasi (2010) found that individuals' epistemic beliefs affect the ways in which they visually inspect different aspects of information. The themes in the works on users' perceptions of reviews are summarized in Table 2.2.

Table 2.2: Works on users' perceptions of reviews.

Themes	References
Factors that predict users' perceptions of reviews	Lim & Van Der Heide (2015) Xu (2014)
Outcomes of users' perceptions of reviews	Casaló et al. (2011) Lim & Van Der Heide (2015)
Human ability to discern review authenticity	George et al. (2014) Toma & Hancock (2012)
Role of epistemic beliefs in discerning review authenticity	Kammerer et al. (2013) Mason & Ariasi (2010)

2.3. Theoretical Model

The works reviewed above point to a lack of connection in the literature. On the one hand, those that distinguish between authentic and fictitious reviews tend to ignore users' perceptions. On the other hand, those that examine users' perceptions of reviews generally overlook if individuals are able to discern authenticity in the first place. It

appears that scholars have yet to disentangle review authenticity by looking at actual differences in tandem with perceived differences between authentic and fictitious reviews.

This calls for investigating how the actual linguistic differences between authentic and fictitious reviews are related to users' perceived review authenticity. As a timely response to the call, this research develops the theoretical model of authentic and fictitious reviews (TAF).

At its crux, the TAF relies on language to distinguish between authentic and fictitious reviews. This is hinged on two reasons. First, accounts intended to describe real experiences have long known to differ linguistically from those concocted out of imagination. This is because the use of language is a potent indicator of the writer's mental dynamics (Johnson & Raye, 1981; Undeutsch, 1967; Vrij et al., 2000). Therefore, authentic reviews that are written with specific post-stay experiences in mind are expected to differ linguistically from fictitious entries that are concocted based on imagination (Ott et al., 2011; Yoo & Gretzel, 2009).

Second, language of reviews is readily available. Hence, it offers a viable way to distinguish between authentic and fictitious reviews. In contrast, metadata such as reviewers' past track record, which could help ascertain authenticity, are not always available in all review websites. For instance, review websites such as Agoda.com and Expedia.com hardly offer any metadata for analysis. Hence, the viability of an approach that relies on metadata to distinguish between authentic and fictitious reviews is confined to specific websites only.

To locate the literature that can form a solid foundation for the TAF, prior works on deception and information processing were extensively reviewed. These works

straddled across several contexts ranging from businesses' financial disclosures (Humphreys et al., 2011), crime settings (Fuller et al., 2011, 2015), and face-to-face interviews (Burgoon & Qin, 2006) to e-commerce websites (Grazioli, 2004; Grazioli & Jarvenpaa, 2003; Xiao & Benbasat, 2011), online dyads (Zhou, Burgoon, Nunamaker et al., 2004), and phishing emails (Vishwanath et al., 2011). They relied on several deception theories and information processing theories that ranged from the information manipulation theory (McCornack, 1992), the reality monitoring theory (Johnson & Raye, 1981), and the self-presentational theory (DePaulo et al., 2003) to the deception detection theory (Johnson et al., 2001), the dual process theory (Deutsch & Gerard, 1955), and the truth bias theory (Vrij & Baxter, 1999). These works together form the conceptual glue to weld the TAF (Figure 2.2), which includes the framework of linguistic cues as well as the framework of user perceptions.

2.3.1. Framework of Linguistic Cues

This part of the TAF, which draws from the deception literature, comprises authenticity as its outcome construct. In addition, it includes four linguistic cues as predictor constructs, namely, comprehensibility, exaggeration, specificity and tentativeness. Each linguistic cue comprises several linguistic traits as shown in Table 2.3.

Authenticity refers to the extent to which an object is a real representation of something else (Grayson & Martinec, 2004). It is commonly conceptualized in terms of originality or genuineness (Kim et al., 2015). For the purpose of this research, authenticity is explicated based on experiences. In other words, the originality of experience is treated as a proxy of authenticity. Hence, as indicated earlier in Section 1.2 (p. 7), authentic

reviews are defined as those contributed by users who had paid, patronized and experienced a given hotel. On the other hand, fictitious reviews refer to those written by participants in a research setting based on imagination without any post-purchase experience of staying in the hotel.

Comprehensibility is a linguistic cue that measures the extent to which a message is easy to read (Berger & Iyengar, 2013; Cukier et al., 2009; Ghose & Ipeiritis, 2011). For the purpose of this research, it refers to the extent to which a review is easy to read. It encompasses three linguistic traits, namely, readability, word familiarity, and structural features.

Specifically, readability is a measure of the amount of effort and expertise required to grasp the meaning of reviews (Zakaluk & Samuels, 1998). Word familiarity is a measure of the degree to which reviews contain words that are easy to recognize (Chall & Dale, 1995). Structural features refer to superficial surface-level characteristics of reviews such as number of characters per word, or number of words per sentence (Cao et al., 2011). Unfamiliar words along with long words and sentences in reviews place more demands on working memory, thereby rendering the entries difficult to process (Graesser et al., 2004). Hence, reviews that are readable, use familiar words, but avoid long words and long sentences are deemed comprehensible (Burgoon et al., 2016; Chall & Dale, 1995; Ghose & Ipeiritis, 2011).

The literature suggests that comprehensibility can be related to authenticity. However, the nature of the relation has been a topic of much debate. Some scholars suggest that authentic reviews could be less comprehensible compared with fictitious ones. This is because individuals writing authentic reviews have lower cognitive load, and

hence, greater bandwidth to craft their sentences (Burgoon & Qin, 2006; Johnson & Raye, 1981; Newman et al., 2003). This in turn makes authentic reviews more sophisticated than fictitious ones. Yet, others expect authentic reviews to be easier to comprehend than fictitious ones. The assumption is that unlike authentic reviews, fictitious ones are written by individuals who are enthusiastic to use sophisticated language in order to sound credible (Ghose & Ipeirotis, 2011; Yoo & Gretzel, 2009). Informed by these works, the following proposition can be made:

P1: Comprehensibility helps ascertain review authenticity.

Exaggeration is a linguistic cue that measures the extent to which a message is fabricated using over-the-top superlatives in order to enhance its intensity (Fletcher et al., 2009; Hu et al., 2012). For the purpose of this research, it refers to the extent to which a review goes overboard to sound convincing. It encompasses four linguistic traits, namely, affect, temporal focus, emphases and syntactic features.

Specifically, affect measures the use of positive or negative emotion words used in reviews to create a lasting impression (Maurer & Schaich, 2011). Temporal focus indicates the chronological spotlight—past, present or future—conveyed in reviews (Tausczik & Pennebaker, 2010). Unlike reviews with a temporal focus on the past, those focusing on the present or the future hints that the experiences described in the entries are likely to persist and recur. Specifically, fictitious reviews could be rich in present and future tense to influence the present and the future revenues of the respective businesses (Dellarocas, 2003; Duan et al., 2008; Gunsch et al., 2000). Emphases refer to the use of hyperbolic expressions in reviews (Yoo & Gretzel, 2009). They serve as rhetorical

devices that elevate the impression of exaggeration (Tausczik & Pennebaker, 2010; Zheng et al., 2006). Syntactic features represent the writing style of reviews at sentence-level. They are commonly captured through the use of punctuations that are often more eloquent than words in exaggerating, and function words that seldom add meaning to content (Afroz et al., 2012; Shojaei et al., 2013). Affective, hyperbolic reviews with a temporal focus on the present or the future, containing several punctuations but few function words are deemed to be exaggerated (Burgoon et al., 2016; Maurer & Schaich, 2011; Tausczik & Pennebaker, 2010).

The literature suggests that exaggeration can be related to authenticity. However, the relation is hardly straightforward. One view holds authentic reviews to be less exaggerated than fictitious ones (Maurer & Schaich, 2011; Yoo & Gretzel, 2009). The former could be innocuous opinion-sharing devices while the latter might go overboard to prove a point. Another view expects fictitious reviews to be less exaggerated than authentic ones. After all, it is possible for fictitious reviews to be written by deviating from orthodox strategies (Godin, 2005). Informed by these works, the following proposition can be made:

P2: Exaggeration helps ascertain review authenticity.

Specificity is a linguistic cue that measures the extent to which a message is rich in details (Bond & DePaulo, 2006). For the purpose of this research, it refers to the extent to which a review is rich in details. It encompasses three linguistic traits, namely, informativeness, perceptual details, and contextual details.

Specifically, informativeness is a measure of the richness of content in reviews (Nakamura, 1991; Ott et al., 2011; Ravid, 2005). Perceptual details indicate the extent to which reviews are rich in words that convey sensory perceptions (Schiffman, 2001). Contextual details indicate the extent to which reviews make references to space and time (Hancock et al., 2005; Johnson & Raye, 1981). Informative reviews that contain perceptual and contextual details are deemed specific (Johnson & Raye, 1981; Ravid, 2005; Rayson et al., 2001).

The literature suggests that specificity can be related to authenticity. However, the nature of the relation lacks clarity. Intuitively, authentic reviews based on real experiences are expected to be more specific than fictitious ones. This is because unlike authentic reviews, fictitious ones lack first-hand experiences (Ott et al., 2011). Nonetheless, given that fictitious reviews are hinged on imagination, some scholars suggests that the lack of experience could be deliberately balanced through concocted specificity (Hancock et al., 2005, 2007). Informed by these works, the following proposition can be made:

P3: Specificity helps ascertain review authenticity.

Tentativeness is a linguistic cue that measures the extent to which a message gives the impression that it has been written with hesitation (Ekman & Friesen, 1969; Latta & Buck, 2008). For the purpose of this research, it refers to the extent to which a review appears to have been written with hesitation. It encompasses three linguistic traits, namely, uncertainty, cognition, and psychological distancing.

Specifically, uncertainty indicates a lack of commitment on the part of the individuals writing reviews toward what is being written (Burgoon et al., 2016). Cognition

indicates the extent of psychical processing involved in writing (Tausczik & Pennebaker, 2010). For example, motion words such as “go” can be strung together easily. Causal words such as “hence” and insight words such as “think” can be used easily to create coherent stories with logical connections (Pennebaker, 1997). However, exclusive words such as “except” are cognitively challenging to write because they differentiate what belongs to a category from what does not (Toma & Hancock, 2012). Psychological distancing refers to the extent to which individuals writing reviews take ownership of the content as conveyed through self-references (Mehrabian, 1967). Reviews rich in uncertainty and words that require low cognitive load but with limited self-references are deemed tentative (Mehrabian, 1967; Pasupathi, 2007; Tausczik & Pennebaker, 2010).

The literature suggests that tentativeness can be related to authenticity. However, the scholarly voices fail to converge. Some scholars argue that authentic reviews would be less tentative compared with fictitious ones. This is because unlike authentic reviews, fictitious entries could be written by individuals feeling the pangs of conscience for using underhanded tactics, or for having their credibility questioned (Vartapetian & Gillam, 2012). Others however suggest that spammers do not necessarily feel guilty in writing fictitious reviews. In fact, they could be enthused by the opportunity to mislead others so easily (Ekman, 1981). Informed by these works, the following proposition can be made:

P4: Tentativeness helps ascertain review authenticity.

Table 2.3: Linguistic cues and linguistic traits.

Linguistic Cues	Linguistic Traits	References
Comprehensibility	Readability	Zakaluk & Samuels (1998)
	Word familiarity	Chall & Dale (1995)
	Structural features	Cao et al. (2011)
Exaggeration	Affect	Maurer & Schaich (2011)
	Temporal focus	Tausczik & Pennebaker (2010)
	Emphases	Yoo & Gretzel (2009)
	Syntactic features	Shojaee et al. (2013)
Specificity	Informativeness	Ott et al. (2011)
	Perceptual details	Schiffman (2001)
	Contextual details	Hancock et al. (2005)
Tentativeness	Uncertainty words	Burgoon et al. (2016)
	Cognitive words	Tausczik & Pennebaker (2010)
	Self-references	Mehrabian (1967)

2.3.2. Framework of User Perceptions

This part of the TAF, which draws from the information processing literature, comprises perceived authenticity as its outcome construct. In addition, it includes four perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness—as predictor constructs coupled with the individual trait of epistemic beliefs as a moderator.

Perceived authenticity refers to individuals’ subjective sense about the originality or genuineness of an object (Blair et al., 2010; Bond & DePaulo, 2006; Kim et al., 2015; Laing et al., 2014). It deals with aspects of individuals’ general disposition about knowledge, and its implications for dealing with evaluative information subjectively regardless of actual authenticity (Lakey et al., 2008; Lehman et al., 2014). For the purpose of this research, perceived authenticity refers to the extent to which users consider reviews to have been written after real experiences.

Among the four perceived linguistic cues, perceived comprehensibility refers to individuals’ subjective sense about the extent to which a message is easy to read (Berger

& Iyengar, 2013; Galante & Thomson, 2017). It is the attribute of a message that helps individuals in sense-making. It should not be misinterpreted as intelligibility, which refers to the extent to which individuals are able to actually understand a message (Derwing & Munro, 2009). For the purpose of this research, perceived comprehensibility refers to the extent to which humans consider reviews as being easy to read. Reviews that are easy to read conceivably aids understanding and sense-making.

Perceived exaggeration refers to individuals' subjective sense about the extent to which a message goes overboard in using over-the-top superlatives in order to enhance its intensity (Chan, 2004; Fletcher et al., 2009; Hu et al., 2012; Maurer & Schaich, 2011). It is the attribute of a message that conveys the extent to which the writer seems desperate to appear convincing (Morris & Lewis, 2010). For the purpose of this research, perceived exaggeration refers to the extent to which humans consider reviews to have gone overboard in an attempt to appear convincing.

Perceived specificity refers to individuals' subjective sense about the extent to which a message is rich in details (Bond & DePaulo, 2006; Ravid, 2005). It is the attribute of a message that makes it appear informative and detailed (Shapiro et al., 1994). For the purpose of this research, perceived specificity refers to the extent to which humans consider reviews as being informative and detailed.

Perceived tentativeness refers to individuals' subjective sense about the extent to which a message is written with hesitation (Ekman & Friesen, 1969; Latta & Buck, 2008). It is the attribute of a message that conveys uncertainty (Latta, 2004). For the purpose of this research, perceived tentativeness refers to the extent to which humans consider a review to be written with hesitation and uncertainty.

Works on the relation between perceived linguistic cues and perceived authenticity offer two competing views. One view suggests that users will be poor in distinguishing between authentic and fictitious reviews. This is because they are inherently inclined to deem any information as being authentic without necessarily paying much attention (Vrij & Baxter, 1999). If users, by default, consider all reviews to be authentic, they would judge authentic reviews correctly yet fail to accurately identify fictitious entries.

In contrast, another view suggests that alertness toward the content of reviews could help users discern review authenticity (Sperber et al., 2010). In particular, works such as Toma and Hancock (2012) suggested that linguistic properties of information inform humans' ability to distinguish between truth and fiction. As the online community becomes increasingly tech-savvy in processing online information, it is perhaps unreasonable to expect users to blindly deem most reviews as being authentic. In fact, they might even consider most reviews as being fictitious (Blair et al., 2010).

Additionally, there are some tacit evidences in the literature that if users are attentive to reviews, they might well manage to distinguish between authentic and fictitious entries with reasonable accuracy. Specifically, if users perceive the actual differences between authentic and fictitious reviews in terms of the linguistic cues of comprehensibility, exaggeration, specificity and tentativeness; they might stand a good chance to discern authenticity of the entries. They might favorably view reviews that are comprehensible (Schrack et al., 2010) and specific (Papathanassis & Knolle, 2011), but with little exaggeration (Connors et al., 2011) and tentativeness (Porter et al., 2012). Hence, the following propositions can be made:

P5: Perceived comprehensibility is related to perceived review authenticity.

P6: Perceived exaggeration is related to perceived review authenticity.

P7: Perceived specificity is related to perceived review authenticity.

P8: Perceived tentativeness is related to perceived review authenticity.

Epistemic beliefs refer to individuals' perceptions of the reliability of knowledge, and the justification for knowing (Hofer, 2004; Hofer & Pintrich, 1997). Epistemic beliefs are intricately related to perceived authenticity, which is known to stem subjectively from aspects of individuals' general disposition about knowledge and knowing (Lakey et al., 2008; Lehman et al., 2014). Specifically, perceptions of the reliability of knowledge include the extent to which knowledge is viewed on a reliable versus unreliable scale. Besides, perceptions of the justification for knowing range from gut feeling to rigorous rules of inquiry (Hofer, 2004; Hofer & Pintrich, 1997; Kammerer et al., 2013).

The literature suggests that epistemic beliefs have the potential to moderate the relation between perceived linguistic cues and perceived authenticity (Bråten et al., 2005; Chiu et al., 2013; Kammerer et al., 2013; Tu et al., 2008). With respect to epistemic beliefs, individuals could be either naïve or robust. Epistemologically naïve individuals are those who perceive knowledge as being relatively reliable, and justify the process of knowing based on gut feeling. They underestimate the challenges involved in evaluating information. On the other hand, epistemologically robust individuals are those who take knowledge with a pinch of salt, and justify the process of knowing through rigorous rules of inquiry. They could be cautious in processing information. Hence, epistemologically robust users could be more careful than epistemologically naïve individuals about the actual differences between authentic and fictitious reviews. In consequence, they could

stand a better chance to discern review authenticity compared with epistemologically naïve individuals. Hence, the following proposition can be made:

P9: Epistemic beliefs moderate the relation between perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity as well as perceived tentativeness—and perceived review authenticity.

Figure 2.2: The theoretical model of authentic and fictitious reviews (TAF).

2.4. Summary of the Literature Review

This chapter has provided a review of the literature related to this research.

Specifically, it began by reviewing the literature on differences between authentic and fictitious reviews. Next, it reviewed related works on users' perceptions of reviews. This was followed by the development of the theoretical model known as the TAF, which is summarized in Table 2.4.

Table 2.4: Summary of the TAF.

	Propositions	Underpinning Literature	
Framework of Linguistic Cues	P1: Comprehensibility helps ascertain review authenticity.	Burgoon & Qin (2006) Newman et al. (2003)	Deception
	P2: Exaggeration helps ascertain review authenticity.	Maurer & Schaich (2011) Tausczik & Pennebaker, 2010	
	P3: Specificity helps ascertain review authenticity.	Hancock et al. (2005, 2007)	
	P4: Tentativeness helps ascertain review authenticity.	Mehrabian (1967) Pasupathi (2007)	
Framework of User Perceptions	P5: Perceived comprehensibility is related to perceived review authenticity.	Blair et al. (2010) Bond & DePaulo (2006) Schrank et al. (2010)	Information Processing
	P6: Perceived exaggeration is related to perceived review authenticity.	Connors et al. (2011) Porter et al. (2012)	
	P7: Perceived specificity is related to perceived review authenticity.		
	P8: Perceived tentativeness is related to perceived review authenticity.		
	P9: Epistemic beliefs moderate the relation between perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity as well as perceived tentativeness—and perceived review authenticity.	Hofer (2004) Kammerer et al. (2013) Lehman et al. (2014)	

Having reviewed the literature, this chapter has laid the foundation for the research to be carried out empirically. Given that the research is divided into two studies—the Linguistic Study, and the User Study—as indicated earlier in Section 1.3 (p. 9), the next chapter is devoted to explaining the methods used to conduct the former.

CHAPTER THREE LINGUISTIC STUDY – METHODS

This chapter elaborates the methods for the Linguistic Study that sought to investigate the research question RQ 1: How do authentic reviews linguistically differ from fictitious entries? The investigation was underpinned by the framework of linguistic cues that drew from the theoretical perspective of deception.

In the Linguistic Study, a ground truth dataset of authentic and fictitious reviews was created. Ground truth refers to a dataset of reviews, the authenticity of which is known a priori with certainty. The data sources included authenticated review websites that facilitated collection of authentic reviews, and hotels that attract large volumes of reviews on such platforms. While authentic reviews were collected from authenticated review websites, fictitious entries were solicited from participants. The identified linguistic cues—comprehensibility, specificity, exaggeration and tentativeness—were operationalized based on the literature. After calculating the operationalized measures for all reviews in the dataset, differences between authentic and fictitious entries were analyzed using classification techniques, followed by feature selection, independent samples t-tests as well as word-level analysis.

As shown in Figure 3.1, the remainder of the chapter is organized into six sections. The first section (p. 38) explains the need for ground truth. The second section (p. 41) justifies the choice of data sources. It covers selection of review websites (p. 42), and identification of hotels (p. 43). The third section (p. 45) explains the data collection procedure, which includes collection of authentic (p. 46), and fictitious reviews (p. 50). The fourth section (p. 59) elucidates the ways in which the linguistic cues in reviews were

operationalized and measured. The fifth section (p. 67) presents the methods for data analysis. Finally, the sixth section (p. 73) summarizes the chapter.

Figure 3.1: Overview of Chapter Three.

3.1. Ground Truth

Ground truth is indispensable for research on information authenticity. It refers to a dataset of messages, the authenticity of which is known a priori with certainty. When what is authentic and what is fictitious is unknown, investigating differences between the two becomes conceivably impossible.

Current approaches to establish ground truth seem to make questionable assumptions (Gokhman et al., 2012). One approach relies on text similarity. For example, works such as Jindal and Liu (2008) labeled unique reviews as authentic, and those that were textually similar to one another as fictitious. However, unique reviews could be fictitious because spammers need not always make blatant copies of existing entries. Conversely, textually similar reviews could be authentic. After all, high text similarity between two reviews might be coincidental.

Another approach relies on trained human annotators. For example, works such as Li et al. (2011) trained a few college students to identify spam. Thereafter, the trained students were asked to annotate reviews as either authentic or fictitious. However, the accuracy of their annotations could not be verified because of the lack of access to the contributors of the entries.

A third approach relies on unauthenticated websites to draw authentic reviews. For example, works such as Ott et al. (2011) ironically drew authentic entries from the unauthenticated website TripAdvisor.com. Given that anyone with Internet access can easily create fake profiles in such unauthenticated websites and submit fictitious reviews, entries from TripAdvisor.com are not necessarily authentic.

As an improvement to the current approaches, this study collects authentic reviews from authenticated websites that allow submission of entries only after valid purchases and stays in a given hotel. On the other hand, fictitious reviews were solicited from participants who had no experience of staying in the hotel. In so doing, the authentic and the fictitious reviews adhere to their respective definitions provided earlier in Section 1.2 (p. 6).

However, this approach to establish ground truth was not free from challenges. It meant that reviewers who contributed authentic reviews could differ from participants who wrote fictitious entries. Such differences would confound the investigation.

For this reason, care was taken to ensure comparable background between reviewers who contributed authentic reviews, and participants who submitted fictitious entries as much as possible in terms of five dimensions. These included country of origin, age, educational profile, travel experiences, and use of review websites.

Reviewers' country of origin could be readily obtained for authentic reviews. Reviewers were grouped into four geographical regions that included America, Asia-Pacific, Europe as well as Middle-East and Africa. Such a geographical grouping is widely used in scholarly works (e.g., Aramberri, 2009; Chen, 2001; Singh, 1997; Zeng & Gerritsen, 2014) as well as in practice by the United Nations World Tourism Organization (UNWTO, 2013), and the World Economic Forum (Blanke & Chiesa, 2011). Participants were recruited to write fictitious reviews keeping in mind the proportions of authentic entries from each of these regions. The target was to have comparable proportions of reviews from the regions in the corpus of authentic reviews as well as in that of fictitious entries. This helped afford a fair comparison between the two corpora.

In terms of age, most reviews are written by young individuals aged 45 years or below (Gretzel et al., 2007; Ip et al., 2012; Ratchford et al., 2003). Hence, fictitious reviews were solicited from participants whose age ranged from 21 to 45 years.

In terms of educational profile, most reviews are written by educated individuals, especially those who have completed secondary/high school (Gretzel et al., 2007; Ip et al., 2012; Rong et al., 2012). Hence, fictitious reviews were solicited from participants who

were minimally undergraduate students in terms of their educational profiles. In particular, participants included undergraduate students, graduate students, and working adults who minimally had undergraduate degrees. Conceivably, all of them had completed secondary/high school education.

In terms of travel experiences, it is conceivable that reviewers writing authentic reviews experienced travelling. That is why they were given access to write reviews in authenticated review websites (Bjering et al., 2015; Crawford et al., 2015; Ott et al., 2012). Hence, fictitious reviews were solicited from participants who had travel experiences in the previous year.

Finally, reviewers are likely to be well-versed with the use of review websites. Hence, fictitious reviews were solicited from participants who were regular readers or contributors on review websites.

To summarize, participants were recruited to write fictitious reviews on meeting the following selection criteria. They must be between 21 and 45 years of age. They had to be minimally undergraduate students in terms of their educational profiles. They must have had travel experiences in the previous year. They had to be well-versed with the use of review websites. Since gender has often been found to be unrelated to review contribution (Ip et al., 2012; Rong et al., 2012), fictitious reviews were solicited from participants without maintaining specific proportions of males and females.

3.2. Data Sources

Identification of data sources included two tasks. The first involved selecting review websites that facilitated collection of authentic reviews. Specifically, reviews

containing titles as well as descriptions were considered because both play crucial communicative roles (Ascaniis & Gretzel, 2012). Hence, it was necessary to select authenticated review websites that solicit entries comprising both titles and descriptions. The second task involved identifying hotels that attract large volumes of reviews on the selected review websites. This was necessary to ensure an adequate number of entries from which authentic reviews could be admitted into the dataset. These two tasks were accomplished on 28 February 2013.

3.2.1. Review Websites

The top 10 travel websites based on user traffic statistics were short-listed for selection. These included Agoda.com, Booking.com, Expedia.com, Hotels.com, Kayak.com, Priceline.com, Southwest.com, TripAdvisor.com, United.com, and Xe.com (Alexa, 2013). Of these, there were six authenticated review websites: Agoda.com, Booking.com, Expedia.com, Hotels.com, Kayak.com, and Priceline.com.

Among these websites, some required reviews to contain titles and descriptions while others did not allow submission of titles. Since both titles and descriptions were taken into consideration, websites that did not allow submitting titles—Booking.com, Kayak.com, and Priceline.com—were excluded.

The remaining three review websites, namely, Agoda.com, Expedia.com and Hotels.com were selected for data collection. Selecting multiple authenticated review websites to collect authentic entries could enhance representativeness of the corpus (Decrop, 1999). Moreover, these review websites have been widely tapped by the scholarly community (Runfola et al., 2012; Stringam & Gerdes, 2010; Tse, 2013).

All the three authenticated review websites invite entries after verified purchases and stays using a three-step process. First, reviewers need to make a hotel booking through the websites. Second, within a few days after checking out of the hotel, they are invited to write reviews via email promptings that contain an exclusive link to the review submission form. Reviewers need to follow the link to contribute their entries. Third, reviews contributed by reviewers are made available online after they had been vetted manually by the website moderators. Since only those with a valid booking are allowed to contribute, reviews in these websites are considered authentic.

Moreover, all the three websites have comparable ways to group hotels into categories—luxury, budget and mid-range. Hotels are assigned hotel ratings based on the availability of facilities such as spas, valet parking, and jacuzzis. A higher hotel rating indicates a superior hotel category. For the purpose of this study, luxury hotels refer to properties that have been consistently assigned hotel ratings of four stars or higher. These are upscale properties that offer the best facilities among all hotels. Budget hotels refer to properties that have been consistently assigned hotel ratings of two stars or lower. These hotels tend to offer minimal facilities. Mid-range hotels are those that have been consistently assigned hotel ratings of three stars. These hotels are generally not as superior as luxury hotels but are better than budget properties. Hotels with inconsistent hotel categories across the three review websites were not considered.

3.2.2. *Hotels*

Hotels located in Asian tourist destinations were short-listed. This is due to three reasons. First, over the last couple of decades, Asia has been one of the fastest growing

tourism regions in the world (Singh, 1997; UNWTO, 2013). Given such unprecedented growth in the popularity of tourism, hotels in Asia provide a fertile ground to investigate the authenticity of reviews. Second, hotels in Asia are known for their online presence (Au & Akiz, 2009). They tend to offer a wider range of online functionalities with emphasis on interactions with consumers compared with hotels located in other continents. Hence, Asian hotels could attract substantial volume of reviews. Third, Asia comprises a wide population of Internet users. The 2.2 billion users in the world are distributed as follows: 44.8% from Asia, 22.1% from Europe, 12.0% from North America, and the remainder from the rest of the world (Internet World Stats, 2012). Thus, Asia provides an interesting context that sees the confluence of adequate tourism growth as well as ubiquitous internet use. Hence, an understanding of authentic and fictitious reviews for hotels in Asia is essential.

Five Asian tourist destinations were chosen, namely, Bangkok, Hong Kong, Kuala Lumpur, Singapore, and Tokyo. They have been consistently rated among the most popular tourist destinations in Asia from 2008 to 2012 by the United Nations World Tourism Organization. Therefore, reviews for hotels in these locations had the potential to foster high levels of online involvement.

From each tourist destination, three hotels were identified: a luxury hotel, a budget hotel, and a mid-range hotel. Identification of hotels in a given tourist destination involved two steps. First, the lists of hotels in the tourist destination featured in Agoda.com, Expedia.com and Hotels.com were inspected. Second, the three review websites were filtered to select the hotel with the highest number of reviews from each of the three hotel categories—luxury, budget and mid-range.

Given that there were five tourist destinations, identification of three hotels yielded a list of 15 properties altogether (5 tourist destinations x 3 hotels). The identified list of 15 properties comprised five luxury hotels, five budget hotels, and five mid-range hotels. All the identified hotels had attracted about 1,000 reviews cumulatively across the selected websites. The details of the hotels are presented in Appendix A (p. 198).

The validity of categorizing hotels as luxury, budget, or mid-range was ensured by examining the consistency of their website-assigned hotel ratings across the three selected 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.

3.3. Data Collection

A ground truth dataset of 1,800 reviews was collected. It included two corpora: one containing 900 authentic reviews, and the other comprising 900 fictitious reviews. Authentic and fictitious reviews were kept equal in number to allow for a fair analysis (Gokhman et al., 2012; Harris, 2012).

The size of 1,800 reviews was deemed appropriate. It not only exceeded the methodological requirements for employing classification, feature selection, independent samples t-tests, and word-level analysis but was also larger than that used in related works such as Yoo and Gretzel (2009), Ott et al. (2011), Harris (2012) and Ong et al. (2014). These works used datasets containing some 80, 800, 800 and 160 reviews respectively.

Both authentic and fictitious reviews were evenly distributed across the 15 identified hotels. Specifically, the corpus of authentic reviews, and that of fictitious

reviews contained 60 entries per hotel (20 positive + 20 negative + 20 moderate). The uniform spread across the three polarities—positive, negative as well as moderate—in the corpora of authentic and fictitious reviews ensured a well-balanced dataset. Authentic reviews were collected in the last week of March 2013 while fictitious entries were solicited from participants via email during April to December 2013. The detailed data collection procedure, the overview of which is depicted in Figure 3.2, is described below.

Figure 3.2: Overview of the data collection procedure.

3.3.1. Authentic Reviews

Authentic reviews were collected from Agoda.com, Expedia.com and Hotels.com in evaluation of the 15 identified hotels. For each hotel, 60 entries (20 positive + 20 negative + 20 moderate) were collected to yield 900 authentic reviews altogether (60

reviews x 15 hotels). Given that all the hotels had attracted huge volumes of authentic reviews as shown in Appendix A (p. 198), there were ample submissions to collect 20 entries of each of the three review polarities.

Reviews were admitted into the corpus based on five inclusion criteria. First, they had to be posted as recently as possible. This was necessary to ensure a fair comparison between authentic and fictitious reviews. Second, both titles and descriptions of reviews must have been written in English. As indicated earlier in Section 1.5 (p. 14), examining non-English reviews is beyond the scope of this research. Third, descriptions of reviews had to be at least more than 150 characters in length. Those shorter than 150 characters offer little room for a linguistic analysis (Harris, 2012; Ott et al., 2011). Fourth, reviews must be meaningful. It is possible for reviewers to write unmistakably irrelevant reviews. Such entries would not have facilitated a sound investigation. Hence, meaningfulness of all titles and descriptions were inspected manually to ensure that they contained evaluation of hotels without any nonsensical texts created through random keystrokes. Fifth, reviews had to be accompanied by metadata about the respective reviewers such as country of origin. This provided the basis to solicit comparable number of fictitious reviews against the volume of authentic entries posted by reviewers across the four regions, namely, America, Asia-Pacific, Europe as well as Middle-East and Africa.

The polarities of reviews—positive, negative and moderate—were determined based on their ratings (Gerdes et al., 2008; Li et al., 2013; O’Mahony & Smyth, 2009). 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 & Lurie, 2013). However, Agoda.com employs a 10-point rating scale. Scales that differ from one another in terms of ranges cannot be linearly interpolated (Dawes, 2008; Ghiselli, 1939; Johnson et al., 1982). In other words, a score of 1 on a 5-point scale is not necessarily equivalent to 2 on a 10-point scale. Rating scales with more options result in higher scores (Ghiselli, 1939). To make ratings from Agoda.com comparable with those of Expedia.com and Hotels.com, the rescaling approach proposed by Dawes (2002) was followed. The scheme adopted to assign polarities to reviews based on their ratings is shown in Table 3.1.

Table 3.1: Scheme to assign review polarities.

Review Websites¹	Review Polarities Based on Ratings		
	Positive	Negative	Moderate
Agoda.com	7.75 or above	5.50 or below	5.51 to 7.74
Expedia.com	4 or 5	1 or 2	3
Hotels.com	4 or 5	1 or 2	3

Note. ¹ Agoda.com uses a 10-point scale whereas Expedia.com and Hotels.com use a 5-point scale.

Collecting 60 authentic reviews for a given hotel involved three steps (cf. Figure 3.3). First, reviews for the hotel available in Agoda.com, Expedia.com and Hotels.com were sorted based on their date of posting, most recent entries being at the top. Second, all reviews across the three websites were inspected chronologically to check if they met the inclusion criteria. Third, when a review met all the criteria, it was admitted into the corpus, and its polarity was noted. If any of the criteria were violated, the review was ignored. The last two steps were repeated until the corpus grew to contain 20 positive, 20 negative, and 20 moderate reviews for the hotel.

These steps were iterated for all the 15 identified hotels to yield the corpus of 900 authentic reviews (300 positive + 300 negative + 300 moderate). In particular, 799 reviews were collected from Agoda.com, 45 from Expedia.com, and 56 from Hotels.com. The reviews were contributed by reviewers from more than some 50 countries of origin. In particular, 71 of them were contributed by reviewers from America, 730 by those from Asia Pacific, 88 by those from Europe, and 11 by those from Middle East and Africa. The reviews were posted during the period ranging from April, 2009 to March, 2013. Specifically, four reviews were posted in 2009, 14 in 2010, 24 in 2011, 320 in 2012, and 538 in 2013.

Figure 3.3: Process of collecting authentic reviews for a given hotel.

Finally, the validity of the review polarities was verified. For this purpose, three research assistants were recruited. They had graduate degrees in Computer Science or

Information Systems with more than one year of professional experience. Moreover, all of them regularly read or contributed reviews, and had travel experience in the previous year.

Each of the three research assistants was randomly assigned one-third of all reviews. They were shown the reviews without their ratings. Hotel names were concealed to avoid biases. They were asked to annotate reviews as either positive, or negative or moderate. No strict deadlines were given to prevent fatigue-induced errors. Cohen's kappa for the agreement between the polarities annotated by the research assistants, and those inferred from the ratings indicated a non-chance level of agreement: $\kappa = 0.84$ (Cohen, 1960). The results of the annotation are shown in Table 3.2.

Table 3.2: Agreement between ratings and annotations for authentic reviews.

		Review Polarities Based on Annotations		
		Negative	Moderate	Positive
Review Polarities Based on Ratings	Negative	268	32	0
	Moderate	7	280	13
	Positive	0	42	258

3.3.2. *Fictitious Reviews*

Fictitious reviews were solicited from participants via email. For each of the 15 hotels, 60 entries (20 positive + 20 negative + 20 moderate) were solicited to yield 900 fictitious reviews altogether (60 reviews x 15 hotels).

Informed by prior works (Ott et al., 2011; Yoo & Gretzel, 2009), participants were instructed to imagine as if they were working for the marketing department of a hotel. Their boss had asked them to write at most six realistic fictitious reviews in English. Each review had to contain a title, and a description of at least 150 characters.

These instructions were meant to aid participants in getting into the groove for writing fictitious reviews, regardless whether malicious intentions were triggered. In any case, as indicated earlier in Section 1.5 (p. 15), disinformative reviews written with malicious intentions and misinformative reviews articulated for non-malicious reasons are collectively treated as fictitious in this research. Furthermore, to adhere to the definition of fictitious reviews as indicated in Section 1.2 (p. 6), participants were asked to submit entry for a hotel only if they had not stayed there earlier.

Initially, a pilot study was conducted with six participants who had graduate degrees in Information Systems, and were regular readers or contributors of reviews. Two of them were instructed to write one positive review each for the six selected hotels from Bangkok and Singapore. The other two were asked to write one negative review each for the same hotels. The remaining two were instructed to write one moderate review each for the same hotels.

The participants of the pilot study were requested to comment on two aspects, namely, the clarity of the instructions, and the perceived difficulty in accomplishing the task. With respect to the former, they unanimously agreed that the instructions were clear. With respect to the latter, the consensus was that participants might take substantial amount of time to write realistic fictitious reviews.

Based on the participants' feedback, two revisions were made. First, the minimum length criterion of "150 characters" in the original instructions was changed to "30 words" in the revised version, assuming five characters per word on average (MacKenzie & Zhang, 1999). This revision was necessary because one participant pointed that the

criterion of “150 characters” could be misinterpreted as “150 words.” After all, individuals engaged in writing tasks relate with words more readily than with characters.

Second, an additional line was inserted in the revised instructions to indicate the estimated time the task could take assuming eight minutes per review on average (Ott et al., 2011). This revision was necessary because one participant indicated that one would “need more time to think and write [fictitious reviews].” Hence, this additional line serves to remind participants that writing fictitious reviews is time-consuming. Additionally, no strict deadline was imposed to ensure the quality of fictitious reviews. If participants were given deadlines, they would have written fictitious reviews perfunctorily and spontaneously. Writing fictitious reviews in such a manner could have resulted in greater cognitive load among participants, thereby providing more linguistic cues for detection (Bond & DePaulo, 2006; DePaulo et al., 2003; Vrij, 2008). A sample email containing instructions as revised based on the pilot study is provided in Appendix B (p. 200).

Collecting fictitious reviews involved four steps (cf. Figure 3.4). In the first step, invitation for voluntary participation in the study was disseminated using a combination of snowball sampling and maximum variation sampling. The former “identifies cases of interest from people who know people who know what cases are information rich” (Creswell, 2012, p. 158), while the latter consists of “determining in advance some criteria that differentiate [the participants]” (Creswell, 2012, pp. 156-157). When differences are maximized before snowballing, better generalization is obtained. Although non-probabilistic, such a purposive sampling approach—selecting participants who serve a specific purpose consistent with the objective of the research (Collingridge & Gantt,

2008)—is considered relevant in research on electronic word-of-mouth and online reviews (Jeong & Koo, 2015; Keller & Berry, 2003).

Based on the researcher's personal contacts, individuals who met the selection criteria of age, education, travel experience and review website familiarity summarized earlier in Section 3.1 (p. 41) were invited for voluntary participation. The invitation was disseminated via email, social networking sites and word-of-mouth. Using email, about 100 seed contacts could be identified. Using social networking sites, about another 150 seed contacts could be reached. More about 50 seed contacts could be contacted through word-of-mouth. All the seed contacts were asked to participate in the study. They were also requested to share the study invitation with as many contacts as possible.

As a part of maximal variation sampling, seed contacts were recruited keeping in mind the proportions of authentic reviews obtained from the four major geographical regions, namely, America, Asia-Pacific, Europe as well as Middle-East and Africa. This helped ensure that the proportions of contributors' country of origin in the corpus of fictitious reviews were comparable to those in the corpus of authentic entries.

If interested to participate, individuals were asked to send an email to the researcher for further instructions. Over 400 participants indicated interest. The response was generally better from seed contacts who were contacted via email or word-of-mouth vis-à-vis those who were reached via social networking sites.

In the second step of collecting fictitious reviews, interested participants were sent emails containing detailed instructions to write fictitious reviews (e.g., Appendix B, p. 200). Each interested participant could choose to write at most six fictitious reviews for

six different hotels. This ensured that the corpus of fictitious reviews was not skewed with disproportionately more entries from any participants.

The participants were randomly assigned to hotels and review polarities. However, they were not tasked to write multiple reviews with different polarities to alleviate sudden context switches during the writing task.

After the instructions were sent to participants, no strict deadlines were imposed to return their fictitious reviews. Nonetheless, callbacks and email follow-ups were used at about weekly intervals up to a maximum of four times as reminders to increase the likelihood of obtaining responses. Not more than four reminders were given because pushing participants excessively might have enticed them to write fictitious reviews perfunctorily. In that case, their fictitious reviews could have provided more linguistic cues for detection (Bond & DePaulo, 2006; DePaulo et al., 2003; Vrij, 2008), thereby hindering a meaningful analysis.

It was necessary that participants must not have stayed in the hotel for which they were writing fictitious reviews (cf. Section 1.2, p. 6). Hence, they were requested not to write fictitious reviews for hotels that they had visited earlier. This meant relying on participants' truthfulness about previous hotel stays. There was only one instance in which a participant admitted to have stayed in a hotel for which a fictitious review was solicited. Following the instructions, the participant did not write review for the visited hotel.

In the third step of collecting fictitious reviews, entries submitted by participants were manually inspected if they were in English, had descriptions more than 150 characters, were meaningful, and adhered to the given instructions. All submissions were

found to be meaningful, and written in English. However, 46 reviews were eliminated as they contained descriptions shorter than 150 characters.

Submissions from participants were checked for missing data points. In particular, one participant was found to have written review descriptions without providing review titles. An email follow-up was used in this case to fill the missing data points.

In the fourth step of collecting fictitious reviews, when a review met all the criteria, it was admitted into the corpus, and its polarity was noted. The demographic information of the respective participant was also archived. If any of the criteria were violated, the review was ignored. The growth of the corpus was closely monitored to ensure that the fictitious reviews were evenly spread across the 15 hotels.

These steps were iterated until the corpus grew to contain at least 20 positive, 20 negative, and 20 moderate reviews for each hotel. Eventually, 909 fictitious reviews were obtained from 287 participants. There were two positive and one negative surplus reviews for the mid-range hotel in Singapore, two surplus positive reviews for the mid-range hotel in Hong Kong, and four surplus positive reviews for the budget hotel in Tokyo. Nine reviews for these hotels and review polarities were randomly eliminated.

Figure 3.4: Process of collecting fictitious reviews for a given hotel.

The final corpus of fictitious reviews contained 900 entries (300 positive + 300 negative + 300 moderate) obtained from 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). The reviews were contributed by participants across some 33 countries of origin. In particular, 69 reviews were contributed by 29 participants from America (13 males, 16 females), 732 by 213 participants from Asia-Pacific (121 males, 92 females), 85 by 37 participants from Europe (20 males, 17 females), and 14 by 5 participants from Middle-East and Africa (2 males, 3 females).

Finally, the validity of the review polarities was verified. The procedure was the same as that employed for verifying the polarity of authentic reviews described earlier in Section 3.3.1 (p. 49). Cohen's kappa for the agreement between the polarities annotated by the research assistants, and those provided by the participants who wrote the fictitious

reviews indicated a non-chance level of agreement: $\kappa = 0.98$ (Cohen, 1960). The results of the annotation are shown in Table 3.3.

Table 3.3: Agreement between ratings and annotations for fictitious reviews.

		Review Polarities Based on Annotations		
Review Polarities Based on Ratings		Negative	Moderate	Positive
	Negative	298	2	0
	Moderate	1	297	2
	Positive	0	2	298

The higher agreement in review polarities among fictitious reviews compared with that among authentic entries (cf. Table 3.2, p. 50) is expected. After all, fictitious reviews were written by participants who received specific instructions to write either positive, or negative, or moderate reviews. However, reviewers articulating authentic reviews were not guided by such instructions. Some examples of incorrectly annotated authentic and fictitious reviews are shown in Appendix C (p. 202).

Overall, the dataset of 1,800 reviews contained a corpus of 900 authentic entries, and a corpus of 900 fictitious entries uniformly distributed across the 15 identified hotels (5 luxury + 5 budget + 5 mid-range). For each hotel, there were 60 authentic reviews (20 positive + 20 negative + 20 moderate), and 60 fictitious reviews (20 positive + 20 negative + 20 moderate). Considering the three hotel categories—luxury, budget and mid-range—as well as the three review polarities—positive, negative and moderate—meant that there were nine combinations. For each of the nine hotel category-review polarity combinations, the dataset contained 200 reviews (9 combinations x 200 reviews = 1800), of which, 100 were authentic and the rest were fictitious. The distribution of reviews across the three hotel categories, and the three review polarities is shown in Table 3.4.

Table 3.4: Distribution of reviews across hotel categories and review polarities.

	Luxury Hotels			Budget Hotels			Mid-range Hotels			Total
	Pos	Neg	Mod	Pos	Neg	Mod	Pos	Neg	Mod	
Authentic	100	100	100	100	100	100	100	100	100	900
Fictitious	100	100	100	100	100	100	100	100	100	900
Total	200	200	200	200	200	200	200	200	200	1,800

Note. Pos, Neg and Mod refers to positive, negative and moderate review polarities respectively.

Given that fictitious reviews were written by participants, a final quality-check was conducted. For this purpose, the 1,800 reviews in the dataset were equally divided among the three research assistants for another round of annotation. Each of them received a random set of 600 reviews (300 authentic + 300 fictitious). They were told that some reviews were authentic while others were fictitious. However, taking the cue from prior works (Du et al., 2013), they were kept ignorant of the exact proportions. The research assistants were asked to indicate if a review was authentic or fictitious to the best of their abilities. They were not given any strict deadlines to prevent fatigue-induced coding errors.

The research assistants were found to accurately identify 415 of the 900 authentic reviews (46.11%), and 469 of the 900 fictitious reviews (52.11%) in the dataset. Their performance resembled random guessing, consistent with the literature on human ability to distinguish between truth and fiction (Bond & DePaulo, 2006; Porter & ten Brinke, 2008). Their sub-par performance assures that the fictitious reviews in the dataset were not only well written but also difficult to be distinguished from the authentic ones.

3.4. Operationalization and Measurement

As indicated earlier in Section 2.3.1 (p. 25), the framework of linguistic cues identified four linguistic cues to distinguish between authentic and fictitious reviews, namely, comprehensibility, exaggeration, specificity and tentativeness. Each was further categorized into linguistic traits (Table 2.3, p. 31), which were operationalized into measurable variables. Efforts were made to operationalize them as exhaustively as possible to precisely validate the framework (McLeod & Pan, 2005).

Comprehensibility encompasses three linguistic traits, namely, readability, word familiarity, and structural features. Readability is commonly operationalized in prior works (Courtis, 1987; Hu et al., 2012; Korfiatis et al., 2012) using seven indicators: Flesch-Kincaid Grade Level (Flesch, 1951; Kincaid et al., 1975), Gunning-Fog Index (Gunning, 1952, 1969), Automated-Readability Index (Smith & Senter, 1967), Simple Measure of Gobbledygook (SMOG) (McLaughlin, 1969), Coleman-Liau Index (Coleman & Liau, 1975), Lasbarhets Index (Bjornsson, 1983), and Rate Index (Anderson, 1983). Of these, SMOG was not considered because it is only meaningful when texts are at least 30 sentences long (Ayello, 1993). Reviews are not necessarily long enough to be examined using SMOG. The remaining six indicators were averaged to create a composite variable, which is henceforth referred as the mean readability index. It allowed harnessing the idiosyncratic uniqueness of the different indicators (Ghose & Ipeirotis, 2011). Word familiarity of reviews was operationalized as the fraction of words in reviews that are easily recognized. For this purpose, each word in the reviews was checked against the Dale-Chall list of 3,000 familiar words (Chall & Dale, 1995), which has been widely used to assess word familiarity (Begeny & Greene, 2014; Misra et al., 2013; Svider et al.,

2013). Structural features were operationalized as the number of characters per word, number of words, fraction of words with 10 or more characters (henceforth, referred as long words), and number of words per sentence (Cao et al., 2011; Harris, 2012; Yoo & Gretzel, 2009; Zheng et al., 2006; Zhou, Burgoon, Twitchell et al., 2004).

Exaggeration encompasses four linguistic traits, namely, affect, temporal focus, emphases and syntactic features. Affect was operationalized based on the fraction of positive emotion words, and negative emotion words as well as emotiveness—the ratio of adjectives and adverbs to nouns and verbs (Burgoon et al., 2016; Maurer & Schaich, 2011; Missen & Boughanem, 2009). Temporal focus was operationalized in terms of the fraction of past, present and future tense words used in reviews (Gunsch et al., 2000; Tausczik & Pennebaker, 2010). Emphases were operationalized in terms of the fraction of firm words (e.g., “never”), upper case characters, and references to hotel names (Yoo & Gretzel, 2009; Zheng et al., 2006). Syntactic features were operationalized as the fraction of question marks, exclamation marks, ellipses, emoticons, all punctuations in general (Afroz et al., 2012; Keshtkar & Inkpen, 2012; Kim et al., 2006; Tsur et al., 2010) as well as function words (Tausczik & Pennebaker, 2010).

Specificity encompasses three linguistic traits, namely, informativeness, perceptual details, and contextual details. Informativeness was operationalized based on lexical diversity—ratio of the number of unique words to that of total words—as well as the fraction of eight parts-of-speech (POS) tags, namely, nouns, adjectives, prepositions, articles, conjunctions, verbs, adverbs, and pronouns (Aune et al., 2005; Nakamura, 1991; Ott et al., 2011; Ravid, 2005; Shie, 2010; Shojaee et al., 2013). Informative texts tend to be lexically diverse, and rich in terms of the first four POS yet scanty in terms of the rest

(Biber et al., 1999; Rayson et al., 2001; Tausczik & Pennebaker, 2010). Perceptual details were operationalized based on the fraction of visual (e.g., “see”), aural (e.g., “hear”) and feeling (e.g., “touch”) words (Hall et al., 1987; Hancock et al., 2005; Johnson & Raye, 1981). Contextual details were operationalized based on the fraction of spatial (e.g., “around”) and temporal (e.g., “until”) words (Bond & Lee, 2005; Vrij et al., 2000).

Tentativeness encompasses three linguistic traits, namely, uncertainty, cognition, and psychological distancing. Uncertainty was operationalized as the fraction of modal verbs (e.g., “could”), fillers (e.g., “I mean”), and hedging (e.g., “perhaps”) words (Ebesu & Miller, 1994; Pasupathi, 2007; Tausczik & Pennebaker, 2010). Cognition was operationalized based on the fraction of causal (e.g., “hence”), insight (e.g., “think”), motion (e.g., “go”), and exclusive (e.g., “except”) words (Boals & Klein, 2005; Lee et al., 2009; Newman et al., 2003). Psychological distancing was operationalized as the fraction of first person singular (e.g., “I”) and plural (e.g., “we”) words (Hancock et al., 2007; Mehrabian, 1967; Zhou et al., 2003).

Cumulatively, there were 44 variables (6 for comprehensibility + 15 for exaggeration + 14 for specificity + 9 for tentativeness). Of these, 34 could be derived using the Linguistic Inquiry and Word Count (LIWC2007) tool (Pennebaker et al., 2007). It offers reliable dictionaries with some 1.6 million words to measure several linguistic indicators (Pennebaker & Graybeal, 2001). Its applicability has been demonstrated in works on deception (Newman et al., 2003) and hotel reviews (Ott et al., 2011).

However, the remaining 10 variables (44 - 34) could not be derived from LIWC2007: (1) mean readability index, (2) word familiarity using the Dale-Chall list, (3)

characters per word, (4) long words, (5) nouns, (6) adjectives, (7) upper case characters, (8) hotel names, (9) ellipses, and (10) emoticons.

A Java program was written to compute reviews' mean readability index, word familiarity, number of characters per word, number of long words, and fraction of upper case characters. Fractions of nouns and adjectives were measured with the help of Stanford Parser's POS tagger (Klein & Manning, 2003).

To compute references to hotel names, a second Java program was written to scan reviews and identify the number of times the hotel names were mentioned. The program was customized to look for numerous variations of the name irrespective of case. For example, for the hotel named D'Oriental Inn, the program checked for phrases such as "d'oriental inn" and "doriental inn". Likewise, for another hotel named L'Hotel Nina et Convention Centre, it checked for phrases such as "l'hotel nina" and "l'nina hotel". Furthermore, the outputs of the program were manually verified to eliminate non-references in the form of homonyms. For example, for the hotel named Hotel Metropolitan Tokyo, one of the fictitious reviews contained the phrase "metropolitan city". Here, the word "metropolitan" was not a reference to the name of the hotel. Such instances were manually sieved out.

Another Java program was written to count the occurrences of ellipses and emoticons. For emoticons, reviews were scanned to specifically look for the nine most popularly used ones: (1) tired (:| , (2) smile :-) , (3) laughing :D , (4) sad :-(, (5) worried :-s , (6) surprise :-o , (7) angry x-(, (8) kiss :-* , (9) love :x (Keshtkar, & Inkpen, 2012).

The program was customized to look for numerous possible variations. For example, to

check for the emoticon on laughing, it scanned reviews for variations such as “:d”, “:D”, “:-d” as well as “:-D”.

The set of 44 variables was measured separately for titles and descriptions of all reviews in the dataset. However, two variables, namely, mean readability index and number of words per sentence were not calculated for titles. This is because these two rely on the number of sentences in a text. Titles of reviews do not necessarily contain complete sentences. Thus, the linguistic cues were measured using a set of 86 variables (cf. Table 3.5) as follows: 42 variables for review titles (44 - 2), and all the 44 variables for review descriptions.

The set of 86 variables was further pruned using a two-step process. First, redundant variables were eliminated to ensure parsimony. For this purpose, correlations among all possible pairs of variables were checked. A correlation greater than 0.80 between any subset of variables suggests redundancy, which often leads to multicollinearity (Licht, 1995; O’Brien, 2007). No subset of variables was found to have inter-correlations greater than 0.80.

Second, variables that occurred rarely were removed. This led to the elimination of three variables, namely, ellipses in review titles, emoticons in review titles, and emoticons in review descriptions. In particular, ellipses in review titles appeared only twice in the dataset. Emoticons occurred in titles and descriptions of reviews only five and four times respectively. Such an approach to prune the set of rarely-occurring variables has also been followed in prior works involving linguistic analysis (Newman et al., 2003).

Table 3.5: Operationalization of the linguistic cues.

Linguistic Cues	Linguistic Traits	Variables	References
Comprehensibility	Readability	(1) Mean readability index ^{1 3}	Ghose & Ipeirotis (2011)
	Word familiarity	(2) Dale-Chall words ³	Chall & Dale (1995)
	Structural features	(3) Characters per word ³	Cao et al. (2011)
		(4) Words ²	Harris (2012)
		(5) Long words ³	Zheng et al. (2006)
		(6) Words per sentence ^{1 2}	
Exaggeration	Affect	(7) Positive emotion words ²	Maurer & Schaich (2011)
		(8) Negative emotion words ²	Missen & Boughanem (2009)
		(9) Emotiveness ²	Yoo & Gretzel (2009)
	Temporal focus	(10) Past tense ²	Gunsch et al. (2000)
		(11) Present tense ²	Tausczik & Pennebaker (2010)
		(12) Future tense ²	
	Emphases	(13) Firm words ²	Tausczik & Pennebaker (2010)
		(14) Upper case characters ³	Yoo & Gretzel (2009)
		(15) Hotel names ³	
	Syntactic features	(16) Question marks ²	Afroz et al. (2012)
		(17) Exclamation marks ²	Keshtkar & Inkpen (2012)
		(18) Ellipses ³	Kim et al. (2006)
		(19) Emoticons ³	Tausczik & Pennebaker (2010)
		(20) All punctuations ²	
		(21) Function words ²	
Specificity	Informativeness	(22) Lexical diversity ²	Biber et al. (1999)
		(23) Nouns ⁴	Nakamura (1991)
		(24) Adjectives ⁴	Ott et al. (2011)
		(25) Prepositions ²	Ravid (2005)
		(26) Articles ²	Rayson et al. (2001)
		(27) Conjunctions ²	Shojaee et al. (2013)
		(28) Verbs ²	Tausczik & Pennebaker (2010)
		(29) Adverbs ²	
		(30) Pronouns ²	
	Perceptual details	(31) Visual words ²	Hall et al. (1987)
		(32) Aural words ²	Hancock et al. (2005)
		(33) Feeling words ²	
	Contextual details	(34) Spatial words ²	Bond & Lee (2005)
		(35) Temporal words ²	Vrij et al. (2000)
Tentativeness	Uncertainty	(36) Modal verbs ²	Ebesu & Miller (1994)
		(37) Fillers ²	Pasupathi (2007)
		(38) Hedging words ²	
	Cognition	(39) Causal words ²	Boals & Klein (2005)
		(40) Insight words ²	Lee et al. (2009)
		(41) Motion words ²	Newman et al. (2003)
		(42) Exclusive words ²	
	Psych. distancing	(43) First person singular words ²	Hancock et al. (2007)
		(44) First person plural words ²	Mehrabian (1967)

Note. ¹ represents variables that could be computed only for descriptions, not for titles of reviews.

² represents variables measured by LIWC.

³ represents variables measured by Java programs.

⁴ represents variables measured using Stanford Parser's POS tagger.

After variable pruning, ellipses were measured only for review descriptions. On the other hand, emoticons were eliminated from the set of variables for being inconspicuous in both titles and descriptions of reviews. Hence, the four linguistic cues were measured using a pruned set of 83 variables ($86 - 3$) as follows: 40 variables for review titles, and 43 variables for review descriptions (cf. Table 3.6).

By measuring the variables for review titles, a reviews-by-title variables matrix (1800 reviews x 40 title variables) was obtained. Likewise, by measuring the variables for review descriptions, a reviews-by-description variables matrix (1800 reviews x 43 description variables) was created. These two matrices were horizontally concatenated to form the reviews-by-variables matrix (1800 reviews x 83 variables), which was used for subsequent analysis.

Table 3.6: Pruned set of variables.

Linguistic Cues	Linguistic Traits	Variables
Comprehensibility	Readability	(1) Mean readability index ¹
	Word familiarity	(2) Dale-Chall words
	Structural features	(3) Characters per word (4) Words (5) Long words (6) Words per sentence ¹
Exaggeration	Affect	(7) Positive emotion words (8) Negative emotion words (9) Emotiveness
	Temporal focus	(10) Past tense (11) Present tense (12) Future tense
	Emphases	(13) Firm words (14) Upper case characters (15) Hotel names
Specificity	Informativeness	(16) Question marks (17) Exclamation marks (18) Ellipses ¹ (19) All punctuations (20) Function words
	Perceptual details	(21) Lexical diversity (22) Nouns (23) Adjectives (24) Prepositions (25) Articles (26) Conjunctions (27) Verbs (28) Adverbs (29) Pronouns
	Contextual details	(30) Visual words (31) Aural words (32) Feeling words
	Tentativeness	(33) Spatial words (34) Temporal words
Tentativeness	Uncertainty	(35) Modal verbs (36) Fillers (37) Hedging words
	Cognition	(38) Causal words (39) Insight words (40) Motion words (41) Exclusive words
	Psych. distancing	(42) First person singular words (43) First person plural words

Note. ¹ represents variables that could be computed only for descriptions, not for titles of reviews.

3.5. Data Analysis

Related works have commonly used five classification techniques (Ghose & Ipeirotis, 2011; O'Mahony & Smyth, 2009; Zheng et al., 2006; Zhou, Burgoon, Twitchell et al., 2004). These include C4.5 decision-tree (C4.5), JRip, logistic regression (LogReg), random forest (RF), and support vector machine (SVM).

However, no single classification technique is known to be consistently superior. Therefore, all of these five were combined. Combining multiple classification techniques has been shown to yield good performance (Kimura & Sridhar, 1991; Wang & Zhu, 2014). Specifically, the meta-classification technique of voting was used. It combined the results of the five classification techniques using average probabilities (Ho et al., 1994).

The classification was implemented using the Waikato Environment for Knowledge Analysis (WEKA) tool (Hall et al., 2009). The parameters in WEKA were mostly set to their default values. However, for SVM, polynomial kernel was used.

A five-fold cross-validation was employed. A k-fold cross-validation involves splitting the dataset into k equal subsets or folds. The classifier is trained on (k-1) folds and tested on the remaining one. This process is repeated k times with each of the k folds used once for testing (Afroz et al., 2012). Performance is reported by taking the micro-average of the results from each of the k folds.

Since the dataset comprised 1,800 reviews, a five-fold cross-validation meant splitting the reviews into five subsets, each containing 360 entries (360 reviews x 5 folds = 1800). The five folds were distributed across reviews for the five tourist destinations, each comprising 360 reviews (360 reviews x 5 tourist destinations = 1800). Training was done on four folds containing 1,440 reviews (360 reviews x 4 folds = 1440), and tested on

the remaining fold with 360 reviews (360 reviews x 1 fold = 360). This process was iterated five times so that every fold was tested once as depicted in Figure 3.5.

Figure 3.5: Iterations of the five-fold cross-validation.

For testing each review, there were four possible outcomes as shown in terms of the confusion matrix in Table 3.7. The outcome is called a true positive (TP) if an authentic review is accurately classified as authentic. It is called a false negative (FN) if an authentic review is inaccurately classified as fictitious. Conversely, the outcome is called a true negative (TN) if a fictitious review is accurately classified as fictitious. It is called a false positive (FP) if a fictitious review is inaccurately classified as authentic.

Table 3.7: Confusion matrix of classification possibilities.

	Classified Authentic	Classified Fictitious
Ground Truth Authentic	True positive (TP)	False negative (FN)
Ground Truth Fictitious	False positive (FP)	True negative (TN)

Classification performance was evaluated in terms of six metrics, namely, precision, sensitivity, specificity, accuracy, F₁-measure, and area under receiver operating characteristic curve (AUC). Precision is the ratio of TP classifications to the number of authentic classifications (TP + FP). Sensitivity, also called recall, is the ratio of TP classifications to the number of authentic instances in the ground truth (TP + FN). Specificity is the ratio of TN classifications to the number of fictitious instances in the ground truth (FP + TN). Accuracy is the ratio of correct classifications (TP + TN) to the number of instances (TP + FN + FP + TN). F₁-measure is the balanced harmonic mean of precision and sensitivity. AUC is estimated by taking the FP rate along the horizontal axis, and the TP rate along the vertical axis. For all these metrics, higher values denote better classification performance (Baldi et al., 2000; Fawcett, 2006; Parker, 2013).

Specifically, to address the research objective RO 1.1 (indicated earlier in Figure 1.1, p. 9), the framework of linguistic cues was used to distinguish between authentic and fictitious reviews using classification. The classification performance of the framework was compared with that of two baselines. The first baseline involved psycholinguistic variables as features for classification. Prior works suggest that the 81 psycholinguistic variables reported by LIWC help distinguish between descriptions of authentic reviews and fictitious entries (Ott et al., 2011, 2013). These variables include words related to themes such as emotion, health, family and religion. All these LIWC variables were computed. The resultant reviews-by-variables matrix (1800 reviews x 81 LIWC variables) was used for classification. This baseline is henceforth referred as the LIWC-baseline.

The second baseline involved bigrams as features for classification. Extant literature suggests that n-grams of review descriptions could be useful to classify

authentic and fictitious entries (Jindal & Liu, 2008; Mihalcea & Strapparava, 2009). In particular, bigrams were found to perform the best (Ott et al., 2011). Hence, after lowercasing and stemming (Porter, 1980), length normalized bigrams were computed for all reviews using WEKA (Hall et al., 2009). For computational efficiency, bigrams were retained only if they occurred at least three times in authentic and fictitious reviews (Forman, 2003). A total of 3,878 such unique bigrams were found. The resultant reviews-by-variables matrix (1800 reviews x 3878 bigrams) was used for classification. This baseline is henceforth referred as the Bigrams-baseline.

These baselines were chosen due to three reasons. First, both are known to perform well in distinguishing between authentic and fictitious reviews (Jindal & Liu, 2008; Mihalcea & Strapparava, 2009; Ott et al., 2011). Thus, these baselines facilitate ascertaining if the proposed framework outperforms current approaches.

Second, both the baselines include some variables without robust theoretical underpinnings. For example, the LIWC-baseline contained variables such as words related to religion, health and family. Likewise, the Bigrams-baseline could influence results fortuitously. If the proposed framework outperforms these two baselines, it would demonstrate the possibility to develop classifiers that are not only theoretically significant but also parsimonious.

Third, both the baselines dealt with review descriptions ignoring the role of review titles in distinguishing between authentic and fictitious entries. Hence, if the proposed framework outperforms the baselines, it would indicate that titles of reviews are also useful to distinguish between authentic and fictitious entries.

Thereafter, to address the research objective RO 1.2 (indicated earlier in Figure 1.1, p. 9), the linguistic traits unique to authentic and fictitious reviews were identified based on two feature selection techniques, namely, information gain (IG), and Chi-squared (χ^2). In particular, IG ranks features used in a classification problem in decreasing order of their relative importance based on entropy (O'Mahony & Smyth, 2009). Besides, χ^2 ranks features by measuring divergence from the expected distribution when feature occurrence is independent of the class value (Forman, 2003). For classification with less than about 100 features, IG and χ^2 are known to perform well (Forman, 2003; Yang & Pedersen, 1997). Moreover, related works have widely used these feature selection techniques (Li et al., 2011; O'Mahony & Smyth, 2009). For the purpose of this study, IG and χ^2 were used to filter the set of 83 variables, and identify those that are important to distinguish between authentic and fictitious reviews. Variables with non-zero IG, and non-zero χ^2 values are generally deemed important for classification (Ananthakrishnan et al., 2011; Kumar & Valsala, 2013; Lowden & Robinson, 1999).

When a variable was found to have non-zero IG values as well as non-zero χ^2 values, it was further examined for authentic and fictitious reviews using independent samples t-tests (Box et al., 1978). A t-test is appropriate to compare differences in means between two samples (McEnroe, & Martens, 2001). For the purpose of this study, t-tests helped to disinter the unique linguistic traits of authentic and fictitious reviews.

When t-tests emerged statistically significant ($p < 0.05$), the results were complemented with qualitative insights from the dataset. To tease out qualitative nuances, titles and descriptions of authentic as well as fictitious reviews were delved deeper using a word-level analysis technique called the Difference-Between-Proportions-Method

(DBPM) (Gerdes et al., 2008). It helps identifying words that differ significantly in their occurrences across two types of documents in the dataset. For the purpose of this study, the DBPM was used to identify words that differed significantly in their occurrences across authentic reviews and fictitious reviews.

The DBPM analysis involved three steps. First, the word count frequency (WCF) of each word was computed among authentic reviews ($WCF_{\text{word, authentic}}$) as well as fictitious reviews ($WCF_{\text{word, fictitious}}$). In the case of DBPM, WCF of a given word is defined as the ratio of number of occurrences of the word in a specific type of document to the number of documents of that type in the dataset (Gerdes et al., 2008). Second, the difference in word count frequency across authentic and fictitious reviews was calculated for each word ($WCF_{\text{word, authentic}} - WCF_{\text{word, fictitious}}$). Third, the differences in word count frequencies for all words were standardized into z-scores. For a word, a positive z-score indicates that the word is more abundant in authentic reviews vis-à-vis fictitious entries. Likewise, a negative z-score indicates that the word is less frequent in the former. An absolute z-score greater than 1.96 for a given word suggests that the occurrence of the word among authentic reviews significantly differs from its occurrence among fictitious reviews ($p < 0.05$). These three steps were repeated separately for review titles and review descriptions.

The results of the DBPM analysis are shown later in Section 4.3 (p. 81) to substantiate the linguistic traits unique to authentic and fictitious reviews. Wherever possible, the results pertaining to the unique linguistic traits are further corroborated with excerpts of authentic and manipulative reviews to offer a qualitative treatment of the dataset. The overall analytical procedure is summarized in Figure 3.6.

Figure 3.6: Summary of the analytical procedure for the Linguistic Study.

3.6. Summary of the Linguistic Study – Methods

This chapter has elaborated the methods used to conduct the Linguistic Study. Specifically, it began by highlighting the importance of creating ground truth for research on information authenticity. Next, it introduced the choice of data sources. Thereafter, it described the methods for data collection, operationalization, measurement and analysis.

Reviews were collected for 15 hotels (5 luxury + 5 budget + 5 mid-range) uniformly distributed across five popular tourist destinations in Asia. On the one hand, authentic reviews were collected from three authenticated review websites, namely, Agoda.com, Expedia.com and Hotels.com. On the other hand, fictitious reviews were solicited from participants via email. A dataset of 1,800 reviews was created. It comprised a corpus of 900 authentic reviews, and another corpus of 900 fictitious entries.

The four linguistic cues identified in the framework of linguistic cues—comprehensibility, exaggeration, specificity and tentativeness—were operationalized as exhaustively as possible. They were measured for each review in the dataset using 83 variables. Analysis involved the classification technique of voting. It was ran using average probabilities of the results from five independent classification techniques that included C4.5, JRip, LogReg, RF and SVM. Classification performance of the proposed framework of linguistic cues was compared against that of two baselines—the LIWC-baseline and the Bigrams-baseline—in terms of six performance metrics. These included precision, sensitivity, specificity, accuracy, F_1 -measure, and AUC. This was followed by feature selection using IG and χ^2 , independent samples t-tests as well as word-level analysis using the DBPM.

Having explained the methods of data collection and analyses for the Linguistic Study, this chapter has laid the foundation to present its findings, which is the purpose of the next chapter. It describes the findings of the two research objectives—RO 1.1 and RO 1.2—corresponding to RQ 1.

CHAPTER FOUR LINGUISTIC STUDY – FINDINGS

This chapter presents the findings of the Linguistic Study that sought to answer the research question RQ 1: How do authentic reviews linguistically differ from fictitious entries? The investigation was underpinned by the framework of linguistic cues that drew from the theoretical perspective of deception. It entailed addressing the two research objectives—RO 1.1 and RO 1.2.

The first research objective (RO 1.1) attempted to distinguish between authentic and fictitious reviews based on linguistic cues. To address this objective, the dataset of 1,800 reviews (900 authentic + 900 fictitious) described earlier in Section 3.3 (p. 45) was analyzed using classification by voting. Performance of the framework was compared with that of two baselines, namely, the LIWC-baseline, and the Bigrams-baseline, which have been introduced earlier in Section 3.5 (p. 69).

The second research objective (RO 1.2) attempted to identify linguistic traits that are unique to authentic and fictitious reviews. To address this objective, the variables used to distinguish between authentic and fictitious reviews were filtered based on two feature selection techniques, namely, information gain (IG), and Chi-squared (χ^2). The filtered variables were delved deeper using independent samples t-tests, and the Difference-Between-Proportions-Method (DBPM) word-level analysis technique. To further offer a qualitative treatment of the dataset, excerpts of reviews are presented to corroborate the linguistic traits of authentic and fictitious reviews.

As shown in Figure 4.1, the remainder of the chapter is organized into five sections. The first section (p. 76) summarizes the descriptive statistics of the dataset used

in this study. The second section (p. 80) presents the results of distinguishing authentic from fictitious reviews in order to address RO 1.1. The third section (p. 81) presents the results pertaining to the unique linguistic traits of authentic and fictitious reviews in order to address RO 1.2. Specifically, it teases out unique traits in terms of the four identified linguistic cues, namely, comprehensibility (p. 83), exaggeration (p. 85), specificity (p. 90), and tentativeness (p. 93). The fourth section (p. 96) discusses the results in relation to the literature. Finally, the fifth section (p. 103) summarizes the chapter.

Figure 4.1: Overview of Chapter Four.

4.1. Descriptive Statistics of the Linguistic Dataset

As indicated earlier in Section 3.3 (p. 45), the dataset for the Linguistic Study included 1,800 reviews (900 authentic + 900 fictitious). These reviews were measured in

terms of the four linguistic cues, namely, comprehensibility, exaggeration, specificity and tentativeness.

Table 4.1 presents the descriptive statistics with respect to comprehensibility. The differences between authentic and fictitious reviews seemed conspicuous in terms of structural features (e.g., words per sentence in descriptions).

Table 4.1: Descriptive statistics with respect to comprehensibility.

	Linguistic Traits	Variables	Authentic Reviews (Mean $\pm$ SD)	Fictitious Reviews (Mean $\pm$ SD)
Titles	Word familiarity	Dale-Chall words	0.66 $\pm$ 0.26	0.64 $\pm$ 0.28
	Structural features	Characters per word	5.28 $\pm$ 1.42	5.34 $\pm$ 1.63
		Words	4.88 $\pm$ 2.34	4.44 $\pm$ 2.28
		Long words	0.07 $\pm$ 0.25	0.11 $\pm$ 0.31
Descriptions	Readability	Mean readability index ¹	10.40 $\pm$ 3.82	10.44 $\pm$ 2.95
	Word familiarity	Dale-Chall words	0.75 $\pm$ 0.07	0.75 $\pm$ 0.07
	Structural features	Characters per word	4.33 $\pm$ 0.37	4.39 $\pm$ 0.34
		Words	81.33 $\pm$ 52.23	82.75 $\pm$ 43.41
		Long words	0.04 $\pm$ 0.03	0.05 $\pm$ 0.03
		Words per sentence ¹	14.49 $\pm$ 7.93	13.62 $\pm$ 4.93

Note. ¹ represents variables that were computed only for descriptions, not for titles of reviews.

Table 4.2 presents the descriptive statistics with respect to exaggeration. The differences between authentic and fictitious reviews seemed conspicuous in terms of affect (e.g., positive emotion words in titles), temporal focus (e.g., past tense in descriptions), emphases (e.g., firm words in titles), and syntactic features (e.g., exclamation marks in titles).

Table 4.2: Descriptive statistics with respect to exaggeration.

Linguistic Traits		Variables	Authentic Reviews (Mean $\pm$ SD)	Fictitious Reviews (Mean $\pm$ SD)
Titles	Affect	Positive emotion words	12.28 $\pm$ 14.42	9.36 $\pm$ 13.31
		Negative emotion words	1.72 $\pm$ 7.08	3.42 $\pm$ 9.47
		Emotiveness	4.28 $\pm$ 16.79	7.72 $\pm$ 20.81
	Temporal focus	Past tense	0.58 $\pm$ 2.88	0.47 $\pm$ 2.53
		Present tense	1.36 $\pm$ 4.82	1.87 $\pm$ 5.41
		Future tense	0.05 $\pm$ 0.86	0.37 $\pm$ 2.55
	Emphases	Firm words	0.47 $\pm$ 3.16	1.42 $\pm$ 5.34
		Upper case characters	0.06 $\pm$ 0.08	0.07 $\pm$ 0.10
		Hotel names	0.02 $\pm$ 0.15	0.04 $\pm$ 0.21
	Syntactic features	Question marks	0.01 $\pm$ 0.28	0.30 $\pm$ 3.25
		Exclamation marks	0.89 $\pm$ 4.72	11.47 $\pm$ 29.09
		All punctuations	18.93 $\pm$ 27.84	28.69 $\pm$ 34.30
		Function words	19.83 $\pm$ 17.84	21.60 $\pm$ 19.47
Descriptions	Affect	Positive emotion words	4.48 $\pm$ 3.35	4.70 $\pm$ 3.50
		Negative emotion words	0.96 $\pm$ 1.44	1.26 $\pm$ 1.77
		Emotiveness	0.40 $\pm$ 0.16	0.44 $\pm$ 0.15
	Temporal focus	Past tense	3.90 $\pm$ 3.65	5.13 $\pm$ 3.78
		Present tense	6.64 $\pm$ 3.77	6.20 $\pm$ 3.89
		Future tense	0.72 $\pm$ 1.03	0.58 $\pm$ 0.89
	Emphases	Firm words	0.99 $\pm$ 1.38	1.39 $\pm$ 1.55
		Upper case characters	0.01 $\pm$ 0.06	0.03 $\pm$ 0.03
		Hotel names	0.14 $\pm$ 0.40	0.24 $\pm$ 0.52
	Syntactic features	Question marks	0.07 $\pm$ 0.52	0.04 $\pm$ 0.32
		Exclamation marks	0.34 $\pm$ 1.21	0.65 $\pm$ 1.70
		Ellipses ¹	0.18 $\pm$ 1.13	0.16 $\pm$ 0.89
		All punctuations	13.94 $\pm$ 5.30	13.51 $\pm$ 5.30
		Function words	52.78 $\pm$ 7.21	54.91 $\pm$ 6.11

Note. ¹ represents variables that were computed only for descriptions, not for titles of reviews.

Table 4.3 presents the descriptive statistics with respect to specificity. The differences between authentic and fictitious reviews seemed conspicuous in terms of informativeness (e.g., nouns in descriptions), and contextual details (e.g., spatial words in titles).

Table 4.3: Descriptive statistics with respect to specificity.

Linguistic Traits	Variables	Authentic Reviews		Fictitious Reviews	
		(Mean $\pm$ SD)		(Mean $\pm$ SD)	
Titles	Informativeness	Lexical diversity	1.00 $\pm$ 0.03	1.00 $\pm$ 0.01	
		Nouns	43.35 $\pm$ 22.52	37.21 $\pm$ 25.54	
		Adjectives	27.24 $\pm$ 20.81	26.58 $\pm$ 22.31	
		Prepositions	5.76 $\pm$ 8.71	5.63 $\pm$ 8.65	
		Articles	2.25 $\pm$ 5.55	2.68 $\pm$ 6.22	
		Conjunctions	4.93 $\pm$ 7.96	3.60 $\pm$ 7.09	
		Verbs	2.19 $\pm$ 6.07	2.96 $\pm$ 7.71	
		Adverbs	2.00 $\pm$ 6.08	2.53 $\pm$ 6.72	
		Pronouns	1.28 $\pm$ 5.05	2.13 $\pm$ 6.21	
	Perceptual details	Visual words	0.29 $\pm$ 2.06	0.21 $\pm$ 1.71	
		Aural words	0.13 $\pm$ 1.38	0.34 $\pm$ 2.71	
		Feeling words	0.18 $\pm$ 1.01	0.26 $\pm$ 2.02	
	Contextual details	Spatial words	8.24 $\pm$ 12.75	4.85 $\pm$ 9.29	
		Temporal words	1.79 $\pm$ 6.25	1.73 $\pm$ 6.30	
Descriptions	Informativeness	Lexical diversity	0.76 $\pm$ 0.09	0.75 $\pm$ 0.83	
		Nouns	28.05 $\pm$ 7.03	25.98 $\pm$ 5.55	
		Adjectives	10.73 $\pm$ 4.20	10.99 $\pm$ 3.94	
		Prepositions	12.94 $\pm$ 3.75	12.20 $\pm$ 3.62	
		Articles	8.28 $\pm$ 3.87	9.00 $\pm$ 3.35	
		Conjunctions	6.62 $\pm$ 2.94	6.65 $\pm$ 2.57	
		Verbs	11.90 $\pm$ 3.92	12.59 $\pm$ 3.54	
		Adverbs	5.54 $\pm$ 3.10	5.81 $\pm$ 3.05	
		Pronouns	8.20 $\pm$ 4.62	9.83 $\pm$ 4.47	
	Perceptual details	Visual words	0.48 $\pm$ 0.99	0.63 $\pm$ 1.12	
		Aural words	0.44 $\pm$ 1.04	0.44 $\pm$ 0.90	
		Feeling words	0.43 $\pm$ 1.00	0.49 $\pm$ 0.94	
	Contextual details	Spatial words	10.27 $\pm$ 4.05	9.18 $\pm$ 3.62	
		Temporal words	4.36 $\pm$ 3.19	4.13 $\pm$ 2.86	

Table 4.4 presents the descriptive statistics with respect to tentativeness. The differences between authentic and fictitious reviews seemed conspicuous in terms of uncertainty (e.g., modal verbs in titles), cognition (e.g., exclusive words in descriptions), and psychological distancing (e.g., first person singular words in descriptions).

Table 4.4: Descriptive statistics with respect to tentativeness.

	Linguistic Traits	Variables	Authentic Reviews (Mean $\pm$ SD)	Fictitious Reviews (Mean $\pm$ SD)
Titles	Uncertainty	Modal verbs	0.36 $\pm$ 2.66	1.05 $\pm$ 4.37
		Fillers	0.02 $\pm$ 0.45	0.14 $\pm$ 1.57
		Hedging words	0.47 $\pm$ 3.11	0.76 $\pm$ 3.48
	Cognition	Causal words	0.31 $\pm$ 2.66	0.26 $\pm$ 1.99
		Insight words	0.63 $\pm$ 3.79	0.92 $\pm$ 4.32
		Motion words	0.78 $\pm$ 3.74	0.51 $\pm$ 2.86
		Exclusive words	4.08 $\pm$ 7.92	3.32 $\pm$ 7.69
	Psych. distancing	First person singular words	0.24 $\pm$ 1.74	0.54 $\pm$ 2.82
		First person plural words	0.24 $\pm$ 1.75	0.53 $\pm$ 2.81
Descriptions	Uncertainty	Modal verbs	1.42 $\pm$ 1.55	1.38 $\pm$ 1.49
		Fillers	0.18 $\pm$ 0.54	0.19 $\pm$ 0.55
		Hedging words	1.89 $\pm$ 1.86	2.01 $\pm$ 1.82
	Cognition	Causal words	0.83 $\pm$ 1.31	0.82 $\pm$ 1.12
		Insight words	1.03 $\pm$ 1.35	1.19 $\pm$ 1.45
		Motion words	1.94 $\pm$ 1.70	1.83 $\pm$ 1.63
		Exclusive words	3.31 $\pm$ 2.49	2.79 $\pm$ 2.34
	Psych. distancing	First person singular words	1.87 $\pm$ 2.42	3.00 $\pm$ 2.85
		First person plural words	1.40 $\pm$ 2.31	1.44 $\pm$ 2.31

4.2. Distinguishing Authentic from Fictitious Reviews

Authentic and fictitious reviews were distinguished using the meta-classification technique of voting. The framework of linguistic cues accurately classified 677 of the 900 authentic reviews, and 714 of the 900 fictitious reviews. The classification performance based on the individual classification techniques is available in Appendix D (p. 203).

The performance of the proposed framework was evaluated in terms of six commonly used metrics, namely, precision, sensitivity, specificity, accuracy, F_1 -measure, and area under receiver operating characteristic curve (AUC). These metrics were also used to examine the performance of the LIWC-baseline, and the Bigrams-baseline, both of which are known to perform well in distinguishing between authentic and fictitious reviews (Jindal & Liu, 2008; Mihalcea & Strapparava, 2009; Ott et al., 2011). Table 4.5 compares the performance of the proposed framework with that of the baselines.

Table 4.5: Comparing the framework of linguistic cues with the baselines.

Approach	Precision	Sensitivity	Specificity	Accuracy	F₁	AUC
LIWC-baseline	0.713	0.704	0.717	71.06 %	0.708	0.767
Bigrams-baseline	0.773	0.714	0.790	75.22 %	0.742	0.839
Framework of linguistic cues	0.784	0.752	0.793	77.28 %	0.768	0.853

The framework of linguistic cues showed generally promising results. It outperformed both the baselines in terms of precision, sensitivity, specificity, accuracy, F₁-measure, and AUC. Overall, the results support the potential of the framework of linguistic cues to distinguish between authentic and fictitious reviews.

4.3. Unique Linguistic Traits

To identify linguistic traits unique to authentic and fictitious reviews, the 83 variables used in the classification were tested using two feature selection techniques, namely, IG and χ^2 . Feature selection was used to identify variables that helped distinguish between authentic and fictitious reviews. Specifically, IG and χ^2 were chosen because both are known to perform well for classification with less than about 100 features (Forman, 2003; Yang & Pedersen, 1997).

Variables with non-zero IG, and non-zero χ^2 values are generally deemed as important features for classification (Ananthakrishnan et al., 2011; Kumar & Valsala, 2013; Lowden & Robinson, 1999). In particular, 41 of the 83 variables had non-zero IG values as well as non-zero χ^2 values as shown in Table 4.6. Of these feature-selected variables, 21 were related to linguistic cues of review titles, and the rest were related to those of review descriptions.

Table 4.6: Information gain and Chi-squared feature selection.

Linguistic cues	Linguistic Traits	Variables	Titles		Descriptions	
			IG	χ^2	IG	χ^2
Comprehensibility	Readability	Mean readability index ¹		NA	0.009	18.407
	Word familiarity	Dale-Chall words	0	0	0	0
	Structural features	Characters per word	0	0	0.012	27.883
		Words	0.020	50.578	0	0
		Long words	0.004	9.203	0.010	25.233
		Words per sentence ¹		NA	0.013	27.920
Exaggeration	Affect	Positive emotion words	0.015	37.267	0	0
		Negative emotion words	0.015	37.143	0.010	24.224
		Emotiveness	0.020	44.565	0.010	24.032
	Temporal focus	Past tense	0	0	0.036	81.973
		Present tense	0	0	0	0
		Future tense	0.007	15.373	0.005	12.027
	Emphases	Firm words	0.010	25.186	0.021	52.074
		Upper case characters	0.020	50.000	0.068	151.280
		Hotel names	0.003	6.897	0.009	22.636
	Syntactic features	Question marks	0.005	9.045	0	0
		Exclamation marks	0.093	200.547	0.008	19.112
		Ellipses ¹		NA	0	0
		All punctuations	0.107	234.224	0	0
		Function words	0.046	104.965	0.016	38.755
Specificity	Informativeness	Lexical diversity	0	0	0	0
		Nouns	0.028	67.980	0.019	48.134
		Adjectives	0	0	0	0
		Prepositions	0	0	0	0
		Articles	0.051	103.332	0.014	35.736
		Conjunctions	0.008	20.847	0	0
		Verbs	0	0	0.008	21.191
		Adverbs	0	0	0	0
		Pronouns	0.009	22.361	0.026	64.405
	Perceptual details	Visual words	0	0	0	0
		Aural words	0	0	0	0
		Feeling words	0	0	0	0
	Contextual details	Spatial words	0.016	39.583	0.013	32.189
		Temporal words	0	0	0	0
Tentativeness	Uncertainty	Modal verbs	0.008	19.663	0	0
		Fillers	0.004	7.027	0	0
		Hedging words	0	0	0	0
	Cognition	Causal words	0	0	0	0
		Insight words	0	0	0	0
		Motion words	0	0	0	0
		Exclusive words	0	0	0.012	30.291
	Psych. distancing	First person singular words	0.014	32.466	0.035	86.309
		First person plural words	0	0	0	0

Note. ¹ represents variables that were computed only for review descriptions. The corresponding IG and χ^2 values for review titles are therefore denoted as Not Applicable (NA).

To delve deeper, the 41 feature-selected variables were further analyzed using a series of t-tests. The results indicated that 38 variables significantly differed between authentic and fictitious reviews. Of these, 19 were related to linguistic cues of titles, and the rest were related to those of descriptions. These variables were deemed to constitute the unique linguistic traits of authentic and fictitious reviews. On the other hand, the other three non-significant variables included mean readability index of review descriptions, articles in review titles, and upper case characters in review titles. The unique linguistic traits of authentic and fictitious reviews in terms of these 38 variables straddling across the four linguistic cues—comprehensibility, exaggeration, specificity and tentativeness—are explained as follows.

4.3.1. Comprehensibility

Titles of authentic and fictitious reviews significantly differed in terms of two structural features, namely, number of words, and fraction of long words. First, number of words in titles was significantly greater for authentic reviews (4.88 ± 2.34) vis-à-vis those for fictitious entries (4.44 ± 2.28); $t(1798) = 4.00$, $p < 0.001$. Since prior studies found that review titles contain about four words on average (Ascaniis, & Gretzel, 2012), the number of entries with titles more than four words were checked. It was found that 469 of the 900 authentic reviews included titles more than four words, the highest being 25 words. On the other hand, 429 of the 900 fictitious reviews contained titles more than four words, the highest being 18 words. Second, long words were significantly scantier in titles of authentic reviews (0.07 ± 0.25) compared with those of fictitious ones (0.11 ± 0.31);

$t(1719.98) = -3.04, p < 0.01$. As indicated earlier in Section 3.4 (p. 60), long words refer to those with 10 or more characters. The DBPM analysis identified only one long word, “experience” ($z = -2.75$), that was significantly scantier in the former. Titles of fictitious reviews contained long words such as “accommodation”, “atmosphere”, and “flexibility”.

Besides, descriptions of authentic and fictitious reviews significantly differed in terms of three structural features, namely, number of characters per word, fraction of long words, and number of words per sentence. First, number of characters per word was significantly lower for descriptions of authentic reviews (4.33 ± 0.37) vis-à-vis those of fictitious entries (4.39 ± 0.34); $t(1785.48) = -2.93, p < 0.01$. Second, long words were significantly scantier in descriptions of authentic reviews (0.04 ± 0.03) compared with those of fictitious ones (0.05 ± 0.03); $t(1798) = -5.30, p < 0.001$. The DBPM analysis identified two long words, namely, “experience” ($z = -6.32$), and “facilities” ($z = -6.09$), which were significantly scantier in the former. Descriptions of fictitious reviews were found to contain other long words such as “recreational”, “conferencing”, and “accessibility”. Third, number of words per sentence was significantly higher in descriptions of authentic reviews (14.49 ± 7.93) than those of fictitious entries (13.62 ± 4.93); $t(1503.32) = 2.80, p < 0.05$. The longest sentence among those in descriptions of authentic reviews contained 103 words. On the other hand, the longest sentence among those in descriptions of fictitious review contained 72 words. The linguistic traits of authentic and fictitious reviews based on comprehensibility are summarized in Table 4.7. These results suggest that comprehensibility helps ascertain review authenticity, thereby lending support to the proposition P1 (cf. Table 2.4, p. 36).

Table 4.7: Linguistic differences based on comprehensibility.

	Linguistic Traits	Variables	t-Stat	Authentic Reviews	Fictitious Reviews
Titles	Structural features	Words	4.00***	Abundant	Scanty
		Long words	-3.04**	Scanty	Abundant
Descriptions	Structural features	Characters per word	-2.93**	Scanty	Abundant
		Long words	-5.30***	Scanty	Abundant
		Words per sentence	2.80*	High	Low

Note. * represents statistically significant result at the $p < 0.05$ level.

** represents statistically significant result at the $p < 0.01$ level.

*** represents statistically significant result at the $p < 0.001$ level.

4.3.2. Exaggeration

Titles of authentic and fictitious reviews significantly differed in terms of all the three variables measuring affect, namely, positive emotion words, negative emotion words, and emotiveness. First, positive emotion words were significantly more abundant in titles of authentic reviews (12.28 ± 14.42) than in those of fictitious ones (9.36 ± 13.31); $t(1798) = 4.46$, $p < 0.001$. The DBPM analysis revealed that titles of authentic reviews used significantly more positive emotion words such as “lovely” ($z = 2.49$). The title of an authentic review containing positive emotion words stated, “Nice hotel with good location.” Second, negative emotion words were significantly scantier in titles of authentic reviews (1.72 ± 7.08) compared with those of fictitious entries (3.42 ± 9.47); $t(1665.32) = -4.29$, $p < 0.001$. Based on the DBPM analysis, negative emotion words such as “bad” were significantly scantier in the former ($z = -2.25$). The title of a fictitious review rich in negative emotion words stated, “disappointing poor service bad staff.” Third, emotiveness was significantly lower in titles of authentic reviews (4.28 ± 16.79) than in those of fictitious ones (7.72 ± 20.81); $t(1720.91) = -3.86$, $p < 0.001$. As indicated earlier in Section 3.4 (p. 60), emotiveness is enhanced by the presence of adjectives or adverbs. The DBPM analysis indicated that adjectives such as “bad” were significantly

scantier in titles of authentic reviews compared with those of fictitious ones ($z = -2.25$). However, it could not identify any specific adverb that was significantly fewer in the former. An emotive title of a fictitious review stated, “Pathetically nice or nicely pathetic.”

In terms of temporal focus, titles of authentic and fictitious reviews significantly differed in the use of future tense. In particular, future tense was significantly scantier in titles of authentic reviews (0.05 ± 0.86) than in those of fictitious ones (0.37 ± 2.55); $t(1101.33) = -3.48$, $p < 0.05$. The title of a fictitious review rich in future tense stated, “Will be back!” However, the DBPM analysis could not identify any specific word in future tense that was significantly fewer in titles of authentic reviews.

Titles of authentic and fictitious reviews significantly differed in using two forms of emphases, namely, firm words, and hotel names. First, firm words were significantly scantier in titles of authentic reviews (0.47 ± 3.16) compared with those of fictitious ones (1.42 ± 5.34); $t(1458.54) = -4.62$, $p < 0.001$. The title of a fictitious review containing the firm word “perfect” stated, “perfect stay perfect place perfect everything.” Second, hotel names were significantly scantier in titles of authentic reviews (0.02 ± 0.15) vis-à-vis those of fictitious entries (0.04 ± 0.21); $t(1628.06) = -2.63$, $p < 0.05$. The title of a fictitious review containing the name of the reviewed hotel stated, “Bye bye [hotel name]. Will never come back!” The DBPM analysis could not identify any specific firm word or hotel name that was significantly fewer in titles of authentic reviews.

Titles of authentic and fictitious reviews also differed significantly in terms of four variables measuring the use of syntactic features, namely, question marks, exclamation marks, all punctuations, and function words. First, question marks were significantly

scantier in titles of authentic reviews (0.01 ± 0.28) than those of fictitious entries (0.30 ± 3.25); $t(911.03) = -2.70$, $p < 0.05$. Second, exclamation marks were significantly scantier in titles authentic reviews (0.89 ± 4.72) compared with those of fictitious ones (11.47 ± 29.09); $t(946.21) = -10.78$, $p < 0.001$. Third, the overall use of punctuations in titles was significantly lower among authentic reviews (18.93 ± 27.84) vis-à-vis its fictitious counterparts (28.69 ± 34.30); $t(1725.45) = -6.62$, $p < 0.001$. Titles of fictitious reviews rich in punctuations included “Is this a real hostel?” and “Overpriced!!!!” Fourth, function words were significantly scantier in titles of authentic reviews (19.83 ± 17.84) than those of fictitious entries (21.60 ± 19.47); $t(1784.34) = -2.00$, $p < 0.05$. The DBPM analysis indicated that function words such as “not” were significantly scantier in the former ($z = -2.14$). The title of a fictitious review containing function words such as “of” and “the” stated, “One of the best I have ever stayed.”

Besides, descriptions of authentic and fictitious reviews significantly differed in terms of two variables measuring affect, namely, negative emotion words and emotiveness. First, negative emotion words were significantly scantier in descriptions of authentic reviews (0.96 ± 1.44) compared with those of fictitious entries (1.26 ± 1.77); $t(1728.38) = -3.90$, $p < 0.001$. The description of a fictitious review rich in negative emotion words pointed, “. . . had a very bad experience . . . suffered a terrible stomach ache . . . diagnosed with food poisoning.” However, the DBPM analysis could not identify any specific negative emotion word that was significantly fewer in descriptions of authentic reviews. Second, emotiveness was significantly lower in descriptions of authentic reviews (0.40 ± 0.16) than those of fictitious ones (0.44 ± 0.15); $t(1798) = -3.42$, $p < 0.05$. The DBPM analysis indicated that adjectives such as “well” ($z = -3.73$) and

adverbs such as “really” ($z = -1.97$) were significantly scantier in the former. The emotive description of a fictitious review expressed, “The housekeeping service is very good . . . not only good but also super prompt . . . provides special offers for corporate . . . at a very convenient location . . .”

In terms of temporal focus, descriptions of authentic and fictitious reviews significantly differed in the use of two forms of tenses, namely, past tense, and future tense. First, past tense was used significantly less in descriptions of authentic reviews (3.90 ± 3.65) vis-à-vis those of fictitious entries (5.13 ± 3.78); $t(1798) = -7.00$, $p < 0.001$. According to the DBPM analysis, the past tense word “was” appeared significantly fewer in the former ($z = -20.91$). The description of a fictitious review rich in past tense pointed, “. . . we were delighted to relax . . . I was surprised . . . it was good to take a dip in the Jacuzzi . . . It was truly a pleasurable stay!” Second, future tense was used significantly more in descriptions of authentic reviews (0.72 ± 1.03) compared with those of fictitious ones (0.58 ± 0.89); $t(1756.48) = 3.20$, $p < 0.05$. According to the the DBPM analysis, the future tense word “will” was significantly more abundant in the former ($z = 6.82$). The description of an authentic review rich in future tense pointed, “. . . will find . . . will definitely . . . probably will . . .”

Descriptions of authentic and fictitious reviews significantly differed in terms of all the three forms of emphases, namely, firm words, upper case characters, and brand references. First, firm words were significantly scantier in descriptions of authentic reviews (0.99 ± 1.38) compared with those of fictitious ones (1.39 ± 1.55); $t(1775.66) = -5.79$, $p < 0.001$. The description of a fictitious review containing firm words such as “must” and “definitely” pointed, “. . . chicken rice is a must try . . . will definitely stay

here . . .” However, the DBPM analysis could not identify any specific firm word that was significantly fewer in descriptions of authentic reviews. Second, upper case characters were significantly scantier in descriptions of authentic reviews (0.01 ± 0.06) than those of fictitious entries (0.03 ± 0.03); $t(1443.88) = -2.14$, $p < 0.05$. The description of a fictitious review with upper case characters pointed, “*WORST 5-STAR HOTEL I HAVE EVER STAYED . . .*” Third, hotel names were significantly scantier in descriptions of authentic reviews (0.14 ± 0.40) than in those of fictitious ones (0.24 ± 0.52); $t(1687.65) = -4.59$, $p < 0.001$. The description of a fictitious review containing the name of the reviewed hotel pointed, “The finest luxury . . . [hotel name] . . . [hotel name] provide many facilities for the customer, starting from . . . in addition [hotel name] also provide . . . All in all, [hotel name] . . .”

Descriptions of authentic and fictitious reviews also differed significantly in terms of two variables measuring the use of syntactic features, namely, exclamation marks, and function words. First, exclamation marks were significantly scantier in descriptions of authentic reviews (0.34 ± 1.21) vis-à-vis those of fictitious entries (0.65 ± 1.70); $t(1630.29) = -4.27$, $p < 0.001$. For example, the description of a fictitious review exclaimed, “. . . Housekeeping services are really noticeable!!! just out of the world!!!” Second, function words were significantly scantier in descriptions of authentic reviews (52.78 ± 7.21) than in those of fictitious entries (54.91 ± 6.11); $t(1750.84) = -6.78$, $p < 0.001$. According to the DBPM analysis, function words such as “*as*” were significantly scantier in the former ($z = -7.16$). The description of a fictitious review containing several function words pointed, “is quite nice, but its location can be problem. It is situated on very crowded place . . . We were there during some congress . . . and it is not prepared . .

.” The linguistic traits of authentic and fictitious reviews based on exaggeration are summarized in Table 4.8. These results suggest that exaggeration helps ascertain review authenticity, thereby lending support to the proposition P2 (cf. Table 2.4, p. 36).

Table 4.8: Linguistic differences based on exaggeration.

	Linguistic Traits	Variables	t-Stat	Authentic Reviews	Fictitious Reviews
Titles	Affect	Positive emotion words	4.46 ^{***}	Abundant	Scanty
		Negative emotion words	-4.29 ^{***}	Scanty	Abundant
		Emotiveness	-3.86 ^{***}	Low	High
	Temporal focus	Future tense	-3.48 [*]	Scanty	Abundant
	Emphases	Firm words	-4.62 ^{***}	Scanty	Abundant
		Hotel names	-2.63 [*]	Scanty	Abundant
	Syntactic features	Question marks	-2.70 [*]	Scanty	Abundant
		Exclamation marks	-10.78 ^{***}	Scanty	Abundant
		All punctuations	-6.62 ^{***}	Scanty	Abundant
		Function words	-2.00 [*]	Scanty	Abundant
Descriptions	Affect	Negative emotion words	-3.90 ^{***}	Scanty	Abundant
		Emotiveness	-3.42 [*]	Low	High
	Temporal focus	Past tense	-7.00 ^{***}	Scanty	Abundant
		Future tense	3.20 [*]	Abundant	Scanty
	Emphases	Firm words	-5.79 ^{***}	Scanty	Abundant
		Upper case characters	-2.14 [*]	Scanty	Abundant
		Hotel names	-4.59 ^{***}	Scanty	Abundant
	Syntactic features	Exclamation marks	-4.27 ^{***}	Scanty	Abundant
		Function words	-6.78 ^{***}	Scanty	Abundant

Note. ^{*} represents statistically significant result at the $p < 0.05$ level.

^{***} represents statistically significant result at the $p < 0.001$ level.

4.3.3. Specificity

Titles of authentic and fictitious reviews significantly differed in terms of three variables measuring informativeness, namely, nouns, conjunctions and pronouns. First, nouns were significantly more abundant in titles of authentic reviews (43.35 ± 22.52) than in those of fictitious entries (37.21 ± 25.54); $t(1770.16) = 5.41$, $p < 0.001$. The DBPM analysis indicated that nouns such as “room” were significantly more abundant in the former ($z = 8.03$). The title of an authentic review rich in nouns stated, “. . . location . . .

thumbs up . . . value for money.” Second, conjunctions were significantly more abundant in titles of authentic reviews (4.93 ± 7.96) compared with those of fictitious entries (3.60 ± 7.09); $t(1775.76) = 3.72$, $p < 0.001$. Based on the DBPM analysis, conjunctions such as “but” were significantly more abundant in the former ($z = 9.34$). The title of an authentic review containing conjunctions stated, “service with smile but small room.” Third, pronouns were significantly scantier in titles of authentic reviews (1.28 ± 5.05) vis-à-vis those of fictitious entries (2.13 ± 6.21); $t(1726.88) = -3.19$, $p < 0.01$. The title of a fictitious review containing pronouns stated, “my review of my 3-day stay.” The DBPM analysis could not identify any specific pronoun that was significantly scantier in titles of authentic reviews.

In terms of contextual details, titles of authentic and fictitious reviews significantly differed in the use of spatial words. In particular, spatial words were significantly more abundant in the former (8.24 ± 12.75) than the latter (4.85 ± 9.29); $t(1643.81) = 6.48$, $p < 0.001$. The DBPM analysis revealed that spatial words such as “location” were significantly more abundant in titles of authentic reviews vis-à-vis those of fictitious entries ($z = 21.03$). A title of an authentic review containing spatial words stated, “great location . . . spacious room.”

Besides, descriptions of authentic and fictitious reviews significantly differed in terms of four variables measuring informativeness, namely, nouns, articles, verbs and pronouns. First, nouns were significantly more abundant in descriptions of authentic reviews (28.05 ± 7.03) than in those of fictitious entries (25.98 ± 5.55); $t(1705.71) = 6.96$, $p < 0.001$. Second, articles were significantly scantier in descriptions of authentic reviews (8.28 ± 3.87) than in those of fictitious entries (9.00 ± 3.35); $t(1761.26) = -4.23$, $p <$

0.001. Third, verbs were significantly scantier in descriptions of authentic reviews (11.90 ± 3.92) than in those of fictitious entries (12.59 ± 3.54); $t(1779.94) = -3.85$, $p < 0.001$. Fourth, pronouns were significantly scantier in descriptions of authentic reviews (8.20 ± 4.62) than in those of fictitious ones (9.83 ± 4.47); $t(1798) = -7.58$, $p < 0.001$. Based on the DBPM analysis, nouns such as “airport” ($z = 2.54$) were significantly more abundant in descriptions of authentic reviews. On the other hand, they contained significantly fewer articles such as “a” ($z = -20.15$), verbs such as “has” ($z = -3.96$), and pronouns such as “my” ($z = -21.83$). The description of an authentic review rich in nouns but with minimal articles, verbs and pronouns illustrated, “*Clean & Tidy after cleaning services . . . Small room with fan . . . Bathroom clean with shampoo, body wash.*”

In terms of contextual details, descriptions of authentic and fictitious reviews significantly differed in the use of spatial words. In particular, spatial words were significantly more abundant in the former (10.27 ± 4.05) than in the latter (9.18 ± 3.62); $t(1774.89) = 5.96$, $p < 0.001$. Based on the DBPM analysis, spatial words such as “near” were significantly higher in descriptions of authentic reviews ($z = 5.44$). The description of an authentic review rich in spatial words pointed, “. . . away from busy street . . . located at the side street, away from the busy main roads . . . quite far . . .” The linguistic traits of authentic and fictitious reviews based on specificity are summarized in Table 4.9. These results suggest that specificity helps ascertain review authenticity, thereby lending support to the proposition P3 (cf. Table 2.4, p. 36).

Table 4.9: Linguistic differences based on specificity.

	Linguistic Traits	Variables	t-Stat	Authentic Reviews	Fictitious Reviews
Titles	Informativeness	Nouns	5.41***	Abundant	Scanty
		Conjunctions	3.72***	Abundant	Scanty
		Pronouns	-3.19**	Scanty	Abundant
	Contextual details	Spatial words	6.48***	Abundant	Scanty
Descriptions	Informativeness	Nouns	6.96***	Abundant	Scanty
		Articles	-4.23***	Scanty	Abundant
		Verbs	-3.85***	Scanty	Abundant
		Pronouns	-7.58***	Scanty	Abundant
	Contextual details	Spatial words	5.96***	Abundant	Scanty

Note. ** represents statistically significant result at the $p < 0.01$ level.

*** represents statistically significant result at the $p < 0.001$ level.

4.3.4. Tentativeness

Titles of authentic and fictitious reviews significantly differed in terms of two variables measuring uncertainty, namely, modal verbs, and fillers. First, modal verbs were significantly scantier in titles of authentic reviews (0.36 ± 2.66) compared with those of fictitious entries (1.05 ± 4.37); $t(1487.09) = -4.03$, $p < 0.001$. The title of a fictitious review containing the modal verb “must” stated, “A must avoid.” Second, fillers were significantly scantier in titles of authentic reviews (0.02 ± 0.45) than in those of fictitious ones (0.14 ± 1.57); $t(1047.56) = -2.18$, $p < 0.05$. The title of a fictitious review containing the filler “like” stated, “Feels like home.” However, the DBPM analysis could not identify any specific modal verb or filler word that appeared significantly fewer in titles of authentic reviews.

In terms of psychological distancing, titles of authentic and fictitious reviews significantly differed in using first person singular words. In particular, first person singular words were significantly scantier in titles of authentic reviews (0.24 ± 1.74) vis-à-vis those of fictitious entries (0.54 ± 2.82); $t(1506.29) = -2.64$, $p < 0.05$. The title of a

fictitious review containing the first person singular word “I” stated, “Worse than I thought.” However, the DBPM analysis could not identify any specific first person singular word that appeared significantly fewer in titles of authentic reviews.

Besides, descriptions of authentic and fictitious reviews significantly differed in cognition as evident from their use of exclusive words. Specifically, exclusive words were significantly more abundant in authentic reviews (3.31 ± 2.49) compared with those of fictitious ones (2.79 ± 2.34); $t(1798) = 4.59$, $p < 0.001$. Based on the DBPM analysis, exclusive words such as “but” were significantly more abundant in the former ($z = 15.06$). The description of an authentic review containing the exclusive word “except” pointed, “This hotel is very nice . . . except . . .”

In terms of psychological distancing, descriptions of authentic and fictitious reviews significantly differed in using first person singular words. In particular, first person singular words were significantly scantier in descriptions of authentic reviews (1.87 ± 2.42) vis-à-vis those of fictitious entries (3.00 ± 2.85); $t(1751.55) = -9.07$, $p < 0.001$. According to the DBPM analysis, the pronoun “I” was significantly scantier in the former ($z = -32.91$). The description of a fictitious review rich in first person singular words pointed, “I am very disappointed . . . I lost my . . . I made a complaint . . . I did not get . . .” The linguistic traits of authentic and fictitious reviews based on tentativeness are summarized in Table 4.10. These results suggest that tentativeness helps ascertain review authenticity, thereby lending support to the proposition P4 (cf. Table 2.4, p. 36).

Table 4.10: Linguistic differences based on tentativeness.

	Linguistic Traits	Variables	t-Stat	Authentic Reviews	Fictitious Reviews
Titles	Uncertainty	Modal verbs	-4.03***	Scanty	Abundant
		Fillers	-2.18*	Scanty	Abundant
	Psych. distancing	First person singular words	-2.64*	Scanty	Abundant
Descriptions	Cognition	Exclusive words	4.59***	Abundant	Scanty
	Psych. distancing	First person singular words	-9.07***	Scanty	Abundant

Note. * represents statistically significant result at the $p < 0.05$ level.

*** represents statistically significant result at the $p < 0.001$ level.

To summarize the results pertaining to the framework of linguistic cues (cf. Table 4.11), authentic and fictitious reviews seemed to exhibit quite disparate linguistic. Moreover, differences identified in titles were not always consistent those in descriptions. Overall, there seems to be no straightforward answer to the question of whether authentic reviews are more comprehensible, exaggerated, specific and tentative vis-à-vis fictitious entries.

Nonetheless, the four linguistic cues—comprehensibility, exaggeration, specificity and tentativeness—cumulatively outperformed the LIWC-baseline and the Bigrams-baseline in ascertaining review authenticity. This confirms that these linguistic cues help ascertain review authenticity, thereby supporting the propositions P1, P2, P3 and P4 in the framework of linguistic cues (cf. Table 2.4, p. 36).

The results further demonstrate that exaggeration had the highest number of variables (19 variables) that helped distinguish between authentic and fictitious reviews followed by specificity (9 variables). However, comprehensibility and tentativeness had fewer such variables (5 variables each). This indicates that exaggeration offered maximal scope to ascertain review authenticity, followed by specificity. Comprehensibility and tentativeness offered relatively less opportunity to distinguish between authentic and

fictitious reviews. Thus, in order to ascertain review authenticity, the linguistic cues could be leveraged in the following order—exaggeration, specificity, followed by comprehensibility or tentativeness.

Table 4.11: Summary of the results pertaining to the framework of linguistic cues.

Propositions	Findings	Additional Insights
P1: Comprehensibility helps ascertain review authenticity.	Supported	Comprehensibility, exaggeration, specificity and tentativeness cumulatively outperformed the LIWC-baseline and the Bigrams-baseline in ascertaining review authenticity. Exaggeration offered maximal opportunity to ascertain review authenticity, followed by specificity. Compared with exaggeration and tentativeness, comprehensibility and tentativeness offered relatively less opportunity to ascertain review authenticity.
P2: Exaggeration helps ascertain review authenticity.	Supported	
P3: Specificity helps ascertain review authenticity.	Supported	
P4: Tentativeness helps ascertain review authenticity.	Supported	

4.4. Comparing Linguistic Study Results with Prior Works

Authentic and fictitious reviews appeared quite similar to one another in two ways. First, authentic reviews did not differ from fictitious ones in terms of readability and word familiarity. In this vein, the literature remains largely inconsistent. Works such as Yoo and Gretzel (2009) suggested that authentic reviews could be more comprehensible, and use more familiar words compared with fictitious entries. However, works such as Harris (2012) suggested that the converse could be true. Moreover, Hu et al. (2012) suggested that authentic and fictitious reviews could exhibit different patterns of readability and word familiarity over time. To augment such works, this study offers a different perspective by suggesting that authentic and fictitious reviews might not necessarily differ

from one another in terms of readability as well as word familiarity. This new finding suggests that spammers could be adept enough to blur the lines between authentic and fictitious reviews in terms of these two linguistic variables.

Second, contrary to prior works (Bond & Lee, 2005; Godin, 2005; Schiffman, 2001), authentic reviews did not contain more perceptual and temporal details compared with fictitious entries. Expression of these details often depends on the time lag between users' service encounter, and review submission. A review posted soon after service encounter is richer in perceptual and temporal details than one submitted after a delay. Given that the selected review websites—Agoda.com, Expedia.com and Hotels.com—solicit authentic reviews a few days after the reviewers check out of hotels, the use of such details in the entries was perhaps comparable to that in fictitious reviews.

Such similarities between authentic and fictitious reviews underscore the difficulty in separating the two. This in turn further highlights the significance of this study. Despite such similarities, the proposed framework of linguistic cues showed promising results in distinguishing between authentic and fictitious reviews. It even outperformed two baselines—the LIWC-baseline and the Bigrams-baseline—that are known to perform well in distinguishing between authentic and fictitious reviews (Jindal & Liu, 2008; Mihalcea & Strapparava, 2009; Ott et al., 2011). The differences between authentic and fictitious reviews based on the four linguistic cues identified in the framework, namely, comprehensibility, exaggeration, specificity and tentativeness, are discussed in light of the literature as follows.

In terms of comprehensibility, authentic and fictitious reviews were distinguishable based on structural features. Authentic reviews comprised more verbose

titles compared with those of fictitious entries. This could be due to the higher difficulty in articulating the latter (DePaulo et al., 2003). Consistent with prior works (Afroz et al., 2012; Burgoon et al., 2003), descriptions of authentic reviews included lengthier sentences than those of fictitious entries despite being comparable in terms of number of words. Moreover, titles and descriptions of authentic reviews used fewer long words vis-à-vis fictitious entries. Spammers perhaps use long words to make fictitious reviews appear sophisticated (Ong et al., 2013; Yoo & Gretzel, 2009).

In terms of exaggeration, authentic and fictitious reviews were distinguishable based on affect, temporal focus, emphases as well as syntactic features. First, with respect to affect, titles of authentic reviews contained more positive emotion words but fewer negative emotion words compared with those of fictitious entries. Descriptions of the former also included fewer negative emotion words than those of the latter. In this vein, two conflicting views exist. One view holds that emotion words are used to create a lasting impact in reviews (Maurer & Schaich, 2011; Mukhejee, Liu & Glance, 2012). Therefore, positive reviews could be replete with positive emotion words, negative entries could be fraught with negative emotion words, and moderate comments could contain both in comparable proportions. The other view suggests that the use of emotion words depends on the mindset of the contributors rather than the intended polarity (Newman et al., 2003). Writing fictitious reviews hinged on imagination is cognitively challenging. As a result, they could contain more negative emotion words regardless of the intended polarity of the content (DePaulo et al., 2003; Tausczik & Pennebaker, 2010). The results of this study appear to lend support for the second view by showing that negative emotion

words were more abundant in fictitious reviews across a dataset that comprised positive, negative as well as moderate entries.

Second, with respect to temporal focus, future tense was scantier in titles yet more abundant in descriptions of authentic reviews compared with the fictitious counterparts. Moreover, past tense was scantier in descriptions of authentic reviews vis-à-vis those of fictitious ones. In general, fictitious reviews are expected to be rich in future tense to convey that the experiences described in the entries are likely to recur. On the other hand, authentic reviews are expected to be rich in past tense because they describe past experiences (Gunsch et al., 2000). However, the results indicate that authentic reviews were not necessarily fraught with past tense. In fact, descriptions of fictitious reviews were rich in past tense perhaps because they were written to deliberately mimic authentic entries. This contradicts earlier findings that fictitious information uses fewer past tense verbs compared with authentic information (Knapp et al., 1974; Watson & Ragsdale, 1981). Nonetheless, it supports the premise that authentic reviews are more imperfect than fictitious entries (DePaulo et al., 2003), which could appear overly clinical and engineered.

Third, with respect to emphases, titles as well as descriptions of authentic reviews contained fewer firm words and hotel names compared with those of fictitious entries. Additionally, descriptions of authentic reviews contained fewer upper case characters than those of fictitious entries. Consistent with prior works (Afroz et al., 2012; Kim et al., 2006; Yoo & Gretzel, 2009; Zhou, Burgoon, Twitchell et al., 2004), fictitious reviews seem to go overboard in order to prove a point.

Fourth, with respect to syntactic features, titles as well as descriptions of authentic reviews contained fewer punctuations and function words compared with those of fictitious entries. The limited use of punctuations in authentic reviews is consistent with the literature (Afroz et al., 2012; Kim et al., 2006). Spammers could use punctuations in fictitious reviews to grab the eyeballs. They could abundantly rely on function words to render the fictitious reviews lengthy without needing to add much meaningful content (Newman et al., 2003; Zheng et al., 2006).

In terms of specificity, authentic and fictitious reviews were distinguishable based on informativeness as well as contextual details. With respect to informativeness, titles as well as descriptions of authentic reviews contained more nouns but fewer pronouns compared with those of fictitious entries. Nouns are the most content rich words (Ravid, 2005). Their abundance in authentic reviews suggests that such entries are generally fraught with meaningful content. On the other hand, pronouns are used to express personal experiences. Their abundance in fictitious reviews points to the subjective nature of such entries. Prior works on text genre suggest that informative texts would be lexically diverse. Moreover, they would contain more nouns, adjectives, prepositions and articles but fewer conjunctions, verbs, adverbs and pronouns compared with imaginative texts (Biber et al., 1999; Nakamura, 1991; Ravid, 2005; Rayson et al., 2001). Apart from the use of nouns and pronouns, the other informativeness-related linguistic traits of authentic and fictitious reviews did not match with those of informative and imaginative texts respectively. Prior works suggest that fictitious reviews could repeat arguments as a form of reinforcement to get the message across (Aune et al., 2005; DePaulo et al., 2003; Sporer & Schwandt, 2006). Repetition reduces lexical diversity. However, comparable

lexical diversity among authentic and fictitious reviews suggests that spammers do not use repetitions in fictitious reviews (Godin, 2005; Vartapetian & Gillam, 2012). Overall, the findings suggest that neither are authentic reviews necessarily informative nor are fictitious reviews imaginative. This could be vestige of reviewers' lackadaisical attitude in writing authentic reviews coupled with spammers' deliberate efforts to make fictitious reviews informative.

With respect to contextual details, titles as well as descriptions of authentic reviews were richer in spatial words compared with those of fictitious entries. Consistent with prior works (Friedman & Tucker, 1990; Newman et al., 2003), authentic reviews were rich in spatial words given their experiential nature. On the other hand, fictitious reviews lacked such words due to their imaginative nature.

In terms of tentativeness, authentic and fictitious reviews were distinguishable based on uncertainty, cognition and psychological distancing. With respect to uncertainty, titles of authentic reviews contained fewer modal verbs and fillers compared with those of fictitious entries. Modal verbs and fillers in texts convey a lack of conviction on the part of the author toward what is being written. Their abundance in fictitious reviews suggests that fictitious reviews are not always written with conviction. This could be attributed to the high cognitive challenge involved in writing fictitious reviews (Johnson & Raye, 1981; Newman et al., 2003; Pasupathi, 2007).

With respect to cognition, descriptions of authentic reviews contained more exclusive words compared with those of fictitious entries. Exclusive words in texts seek to differentiate between what belongs to a category and what does not (Toma & Hancock, 2012). In general, such a differentiation is relatively easier to articulate while describing

real experiences than while concocting stories based on imagination (Boals & Klein, 2005; Newman et al., 2003; Tausczik & Pennebaker, 2010). Hence, the finding that authentic reviews contained more exclusion words compared with fictitious entries does not come as a surprise.

With respect to psychological distancing, titles as well as descriptions of authentic reviews contained fewer first person singular words compared with those of fictitious entries. Prior works suggest that individuals engaged in fictitious behavior often feel guilt for misleading others or having their reputation questioned (Ekman, 1981; Vartapetian & Gillam, 2012; Vrij et al., 2000). This in turn prompts them to use fewer first person singular words, and more first person plural words. In this way, they could dissociate themselves from their fictitious content, and transfer accountability to some imaginary groups (Freud, 1914; Mehrabian, 1967; Zhou et al., 2003). However, such a trend was largely inconspicuous. Nonetheless, the richness of first person singular words in fictitious reviews complies with recent studies that tested the use of pronouns specifically in the context of authentic and fictitious reviews (Harris, 2012; Yoo & Gretzel, 2009). A possible explanation is that writing fictitious reviews is cognitively more demanding than articulating authentic entries (Newman et al., 2003). When individuals are engaged in a challenging task, they tend to draw attention toward themselves by using first person singular words abundantly (Rude et al., 2004). Moreover, the use of first-person singular words could be a sign of psychological distress (Bucci & Freedman, 1981; Stirman & Pennebaker, 2001). The linguistic differences between authentic and fictitious reviews are summarized in Table 4.12.

Table 4.12: Linguistic differences between authentic and fictitious reviews.

	Linguistic Cues	Linguistic Traits	Variables (word example where applicable)	Authentic Reviews	Fictitious Reviews
Titles	Comprehensibility	Structural features	Words Long words (“experience”)	Abundant Scanty	Scanty Abundant
	Exaggeration	Affect	Positive emotion words (“lovely”) Negative emotion words (“bad”) Emotiveness	Abundant Scanty Low	Scanty Abundant High
		Temporal focus	Future tense (“will”)	Scanty	Abundant
		Emphases	Firm words (“perfect”) Hotel names	Scanty Scanty	Abundant Abundant
		Syntactic features	Question marks Exclamation marks All punctuations Function words (“not”)	Scanty Scanty Scanty Scanty	Abundant Abundant Abundant Abundant
	Specificity	Informativeness	Nouns (“room”) Conjunctions (“but”) Pronouns (“my”)	Abundant Abundant Scanty	Scanty Scanty Abundant
		Contextual details	Spatial words (“location”)	Abundant	Scanty
	Tentativeness	Uncertainty	Modal verbs (“must”) Fillers (“like”)	Scanty Scanty	Abundant Abundant
		Psych. distancing	First person singular words (“I”)	Scanty	Abundant
Descriptions	Comprehensibility	Structural features	Characters per word Long words (“accessibility”) Words per sentence	Scanty Scanty High	Abundant Abundant Low
	Exaggeration	Affect	Negative emotion words (“terrible”) Emotiveness	Scanty Low	Abundant High
		Temporal focus	Past tense (“was”) Future tense (“will”)	Scanty Abundant	Abundant Scanty
		Emphases	Firm words (“must”) Upper case characters Hotel names	Scanty Scanty Scanty	Abundant Abundant Abundant
		Syntactic features	Exclamation marks Function words (“as”)	Scanty Scanty	Abundant Abundant
	Specificity	Informativeness	Nouns (“airport”) Articles (“a”) Verbs (“has”) Pronouns (“my”)	Abundant Scanty Scanty Scanty	Scanty Abundant Abundant Abundant
		Contextual details	Spatial words (“near”)	Abundant	Scanty
	Tentativeness	Cognition	Exclusive words (“except”)	Abundant	Scanty
		Psych. distancing	First person singular words (“I”)	Scanty	Abundant

4.5. Summary of the Linguistic Study – Findings

This chapter has presented the findings of the Linguistic Study. Specifically, it began by presenting the descriptive statistics of the dataset used for the study. Next, it

presented the results of RO 1.1, which distinguished between authentic and fictitious reviews based on linguistic cues. The results were generally promising. The proposed framework, which encompassed 83 variables, outperformed the two selected baselines—the LIWC-baseline and the Bigrams-baseline—in distinguishing between authentic and fictitious reviews.

Thereafter, this chapter presented the results of RO 1.2, which identified linguistic traits unique to authentic and fictitious reviews. These linguistic traits included 38 of the 83 total variables in the framework of linguistic cues. Moreover, they straddled across all the four linguistic cues identified in the framework, namely, comprehensibility, exaggeration, specificity and tentativeness. The results have been discussed in light of the literature.

Having presented the findings of the Linguistic Study, this chapter has laid the foundation to carry out the User Study. Hence, the next chapter is devoted to explaining the methods used to conduct it.

CHAPTER FIVE USER STUDY – METHODS

This chapter elaborates the methods for the User Study that sought to investigate the research question RQ 2: To what extent are users able to perceive the authenticity of reviews? The investigation was underpinned by the framework of user perceptions that drew from the theoretical perspective of information processing.

In the User Study, an online survey was conducted. Surveys are appropriate to gain information about human opinion or perception. The survey was specifically conducted online to attract participants easily by overcoming time and place constraints (Chang & Chen, 2008). The online nature of the survey enabled participants to remain in their familiar settings during participation instead of having to come to unfamiliar laboratory settings (Reips, 2002).

The survey stimuli comprised a review website presenting a set of reviews. For this purpose, a simulated review website was designed. A subset of authentic and fictitious reviews was selected from the dataset used in the Linguistic Study—described earlier in Section 3.3 (p. 45)—to be presented on the website. Participants' perceptions of reviews—some authentic, and others fictitious—were obtained using a questionnaire that sought closed-ended as well as open-ended responses. Analytical approaches included structural equation modeling and content analysis.

Additionally, as indicated earlier in Section 1.3 (p. 9), a Robustness Check was conducted to confirm the findings of this research. It investigated the possibility of developing a guideline to help humans identify the linguistic differences between authentic and fictitious reviews. This was necessary because guidance on ways to evaluate

information is known to improve individuals' information-processing strategies (Argelagós & Pifarré, 2012). The guideline was developed by relying on the unique linguistic traits of authentic and fictitious reviews identified in the Linguistic Study (cf. Section 4.3, p. 81). After a few rounds of pre-test, the efficacy of the guideline was examined by using it as an intervention in an experimental design.

As shown in Figure 5.1, the remainder of the chapter is organized into six sections. The first section (p. 107) introduces the survey stimuli used in the study. In particular, it describes the simulated review website (p. 108), and the choice of authentic as well as fictitious reviews (p. 110). The second section (p. 112) describes the survey procedure used to examine humans' perceptions of authentic and fictitious reviews. The third section (p. 117) develops the questionnaire that sought both closed-ended and open-ended responses. It presents the operationalized measures (p. 118), and explains the ways they were evaluated (p. 123). The fourth section (p. 126) presents the analytical approaches used. Approaches for both quantitative data analysis (p. 126), and qualitative data analysis (p. 131) are described. The fifth section (p. 134) elaborates the methods used to conduct the Robustness Check. Finally, the sixth section (p. 140) summarizes the chapter.

Figure 5.1: Overview of Chapter Five.

5.1. Survey Stimuli

The survey stimuli comprised a simulated review website LoveToTravel.com, and a set of reviews. Several versions of the website were created and presented to different groups of participants. Moreover, a set of 54 reviews (27 authentic + 27 fictitious) was identified from the dataset used in the Linguistic Study to be displayed on the website. The reviews were uniformly distributed across the nine combinations obtained by crossing

the three hotel categories—luxury, budget and mid-range—with the three review polarities—positive, negative and moderate—as shown in Table 5.1.

Table 5.1: Reviews used as stimuli across hotel categories and review polarities.

Hotel Categories	Review Polarities	# Reviews¹
Luxury	Positive	6 reviews (3 A + 3 F)
	Negative	6 reviews (3 A + 3 F)
	Moderate	6 reviews (3 A + 3 F)
Budget	Positive	6 reviews (3 A + 3 F)
	Negative	6 reviews (3 A + 3 F)
	Moderate	6 reviews (3 A + 3 F)
Mid-range	Positive	6 reviews (3 A + 3 F)
	Negative	6 reviews (3 A + 3 F)
	Moderate	6 reviews (3 A + 3 F)
Total		54 reviews (27 A + 27 F)

Note. ¹ A and F denote authentic and fictitious reviews respectively.

5.1.1. Simulated Review Website

The simulated review website was designed in ways so as to enhance control over extraneous factors such as participants' familiarity and preferences (Boush & Loken, 1991; Chang, 2016; Cook & Campbell, 1976; Keller & Aaker, 1992). After all, individuals might have varying levels of familiarity with different review websites that adopt distinctive interface designs. For example, Expedia.com displays a dedicated link on a hotel's information page to redirect users to the reviews attracted by the property. In contrast, Agoda.com features the reviews directly at the bottom of the hotel's information page. Moreover, individuals could also have preferences for specific hotel brands. Their perceptions of reviews could be biased if they come across their preferred brands. To control for such extraneous factors, the simulated review website was referred by the

imaginary name LoveToTravel.com. It presented reviews for an imaginary property Three2One Hotel.

Additionally, it was also necessary to make the simulated review website realistic (Viswanathan, 2005). For this purpose, three features were included in the website: hyperlinks to other parts of the website, details of the hotel for which reviews were displayed, and a realistic presentation of reviews (Ludwig et al., 2013; Sparks & Browning, 2011). The website showed hyperlinks to other parts of the website such as Home, Hotels and FAQ.

Details of the hotel included a location and a description. The location of Three2One Hotel was indicated as Hong Kong, a popular tourist destination in Asia. Its description contained phrases such as “stylish accommodation” when instantiated as a luxury hotel, those such as “no-frills accommodation” when instantiated as a budget hotel, and those such as “comfortable accommodation” when instantiated as a mid-range hotel.

Consistent with the design of contemporary review websites, ratings and titles of reviews were presented more conspicuously than descriptions (Ludwig et al., 2013). The posting time of reviews was also indicated. To control for recency, all reviews were indicated as having been posted a day ago. The website showed three reviews concurrently without any need to scroll because users read about three entries on average for decision-making (Bambauer-Sachse & Mangold, 2011, 2013). Figure 5.2 shows a screenshot of the design template used for LoveToTravel.com. Examples of screenshots for the actual websites are provided in Appendix E (p. 204).

[HOME](#)
[HOTELS](#)
[FLIGHTS](#)
[RENTALS](#)
[OFFERS](#)
[REGISTER](#)
[LOGIN](#)
[ABOUT US](#)
[CONTACTS](#)
[FAQ](#)

Three2One Hotel
 Insert hotel description

 Posted 1 day ago

Insert first review rating --- "Insert first review title"
 Insert first review description

 Posted 1 day ago

Insert second review rating --- "Insert second review title"
 Insert second review description

 Posted 1 day ago

Insert third review rating --- "Insert third review title"
 Insert third review description

Figure 5.2: Design template used for the simulated review website.

5.1.2. Choice of Reviews

Selecting 54 reviews (27 authentic + 27 fictitious) from the dataset of 1,800 entries (900 authentic + 900 fictitious)—used in the Linguistic Study as described earlier in Section 3.3 (p. 45)—comprised a three-step process. First, the reviews were filtered to identify only the accurately classified authentic reviews, and the accurately classified fictitious reviews based on the framework of linguistic cues. This ensured that the selection of reviews for the User Study was informed by the results of the Linguistic Study. As indicated earlier in Section 4.2 (p. 80), the proposed framework of linguistic cues accurately identified 677 of the 900 authentic reviews, and 714 of the 900 fictitious reviews in the dataset. Put differently, these 677 authentic reviews, and 714 fictitious entries were typical in nature, and largely consistent with the overall results pertaining to the framework of linguistic cues. These 1,391 reviews (677 authentic + 714 fictitious) formed a filtered pool from which reviews were selected as the survey stimuli.

Second, from the filtered pool, reviews with specific location references (e.g., “Nathan Road”), brand references (e.g., specific names of hotels, restaurants, or malls), or cultural references (e.g., “China travellers”) were eliminated. When used as stimuli, such reviews might have confounded the results. This step yielded a smaller pool of 985 reviews (518 authentic + 467 fictitious).

Third, from the smaller pool of 985 reviews, stratified random sampling followed by random selection was employed to identify the final set of 54 reviews (27 authentic + 27 fictitious). The sets of the authentic as well as the fictitious reviews were stratified across the nine combinations of hotel categories, and review polarities (cf. Table 5.1, p. 108). This resulted in a total of 18 strata (9 for authentic + 9 for fictitious). Thereafter, a set of three reviews were randomly selected from each stratum, thereby yielding 54 entries altogether ($18 \text{ strata} \times 3 \text{ reviews} = 54$). The three steps are depicted in Figure 5.3.

Figure 5.3: Three-step process to choose reviews as the stimuli.

5.2. Survey Procedure

This study targeted to recruit about 300 participants for the survey. Each of them would be exposed to a total of three reviews on LoveToTravel.com because users read about three entries on average for decision-making (Bambauer-Sachse & Mangold, 2011). The approach of showing three reviews to participants for evaluation is informed by the literature (Bambauer-Sachse & Mangold, 2011, 2013; Connors et al., 2011). In any case,

showing all the 54 reviews selected as the stimuli to each participant would have been cognitively onerous. It could have resulted in fatigue among participants, thereby impairing the quality of data.

As shown earlier in Table 5.1 (p. 108), there were nine combinations of hotel categories, and review polarities. For each combination, there were six reviews (3 authentic + 3 fictitious), thereby yielding 54 entries altogether. These reviews were split into 18 sets to setup the arrangement of three reviews per participant (18 sets x 3 reviews = 54 reviews). These sets were included on 18 different versions of LoveToTravel.com with three reviews each. Half of these versions contained two authentic reviews, and one fictitious review. The other half contained two fictitious reviews, and one authentic review. To control for order effects (Chang, 2016), the 18 versions of LoveToTravel.com were unique in terms of hotel categories, review polarities as well as sequence of authentic and fictitious reviews. The distribution of the 54 reviews across the 18 versions of LoveToTravel.com is shown in Table 5.2.

Table 5.2: Distribution of reviews across versions of LoveToTravel.com.

Hotel Categories	Review Polarities	# Reviews ¹	Website Versions	Sequence ¹
Luxury	Positive	6 reviews (3A + 3F)	1	AAF
			2	FFA
Luxury	Negative	6 reviews (3A + 3F)	3	AFA
			4	FAF
Luxury	Moderate	6 reviews (3A + 3F)	5	FAA
			6	AFF
Budget	Positive	6 reviews (3A + 3F)	7	FAA
			8	AFF
Budget	Negative	6 reviews (3A + 3F)	9	AAF
			10	FFA
Budget	Moderate	6 reviews (3A + 3F)	11	AFA
			12	FAF
Mid-range	Positive	6 reviews (3A + 3F)	13	AFA
			14	FAF
Mid-range	Negative	6 reviews (3A + 3F)	15	FAA
			16	AFF
Mid-range	Moderate	6 reviews (3A + 3F)	17	AAF
			18	FFA

Note. ¹ A and F denote authentic and fictitious reviews respectively. Three pairs of review sequences were used: AAF-FFA, AFA-FAF, and FAA-AFF. Each pair was used once for each of the three hotel categories—luxury, budget and mid-range—as well as once for each of three review polarities—positive, negative and moderate. This ensured that the 18 versions of LoveToTravel.com were unique in terms of hotel categories, review polarities as well as sequence of authentic and fictitious reviews.

Individuals were recruited for participation using snowball sampling, which is appropriate in research on electronic word-of-mouth and online reviews (Jeong & Koo, 2015; Keller & Berry, 2003), on meeting four eligibility criteria. First, they must be between 21 and 45 years of age. After all, young individuals aged below 45 years read reviews regularly (Gretzel et al., 2007; Ip et al., 2012; Ratchford et al., 2003). Second, they had to be minimally undergraduate students in terms of their educational profile. This was necessary because reviews are mostly read by educated individuals, especially those who have completed secondary/high school (Gretzel et al., 2007; Ip et al., 2012; Rong et al., 2012). Third, they must have read reviews prior to making purchases. This ensured that they were appropriate to complete the task at hand. Fourth, they must not have taken

part in any component of this research earlier. This was necessary to control for potential biases.

The invitation for voluntary participation was disseminated through an advertisement that is available in Appendix F (p. 206). The advertisement was posted on several notice boards inside the campus of a large public university. Additionally, it was distributed via email among about 60 contacts of the researcher. All those who participated were requested to share the study invitation with as many contacts as possible. Eventually, 473 participants took part in the survey between December 2014, and May 2015.

The survey procedure comprised five steps. First, each participant was randomly assigned to one of the 18 versions of LoveToTravel.com. Random assignment was necessary to control for participants' individual differences (Park et al., 2007). The progress of the data collection was carefully monitored to ensure that all the versions received comparable number of participants as much as possible. Participants were given the URL of the assigned version of LoveToTravel.com via email.

Second, participants were asked to imagine visiting a hotel review website to evaluate a hotel prior to booking for their forthcoming trip. They were introduced to the assigned version of the simulated review website LoveToTravel.com, and to the imaginary property Three2One Hotel in Hong Kong. Informed by prior works (Papathanassis & Knolle, 2011), participants were instructed to assume that it was a relevant and affordable accommodation choice for their trip. This allowed controlling for participants' preferences and budget constraints.

Third, participants were asked to carefully read the three reviews. To heighten suspicion, they were told that at least one of the reviews was fictitious. Taking the cue from prior works (Du et al., 2013; Reips, 2002), they were requested to pay undivided attention during participation given the difficulty in distinguishing between authentic and fictitious reviews. Participants were promised SGD 5 as a token of appreciation. They were further informed that that they could receive additional SGD 5 if they were able to predict authenticity of all the three reviews accurately. This strategy was informed by the literature as a way to motivate sincerity among participants (Blair et al., 2010; Ekman & Friesen, 1974; George et al., 2014). In reality however, all participants were supposed to receive the maximum possible amount—SGD 10—regardless of their responses.

Fourth, participants were required to answer the questionnaire, which is presented later in Section 5.3 (p. 117). The questionnaire asked participants' perceptions about each of the three reviews. Thereafter, it sought information about participants' individual differences including their epistemic beliefs.

Finally, participants received an incentive of about SGD10 as a token of appreciation. A debriefing statement was displayed at the end of the survey informing that they were promised additional incentives only in order to heighten their vigilance. In reality however, all participants were to receive equal monetary compensation regardless of their responses (Blair et al., 2010; Ekman & Friesen, 1974; George et al., 2014). Participants were requested to share the study invitation with as many contacts as possible. They were reminded not to disclose their participation experience with anyone.

Of the 473 participants who took part in the survey, 93 did not complete. Data obtained from the remaining 380 participants (473 - 93) were admitted for analysis. Table

5.3 shows the distribution of the participants across the 18 versions of LoveToTravel.com. Each version received minimally 16 participants. This compares favorably with similar prior works (Toma & Hancock, 2012).

Table 5.3: Distribution of participants across versions of LoveToTravel.com.

Website Versions	Total Participants	Incomplete Responses	Complete Responses
1	24	6	18
2	21	4	17
3	28	8	20
4	26	5	21
5	19	3	16
6	27	4	23
7	22	5	17
8	27	5	22
9	29	7	22
10	26	5	21
11	31	5	26
12	28	4	24
13	25	5	20
14	29	4	25
15	27	4	23
16	31	4	27
17	31	9	22
18	22	6	16
Total	473	93	380

5.3. Questionnaire Development

As explained earlier in Section 2.3.2 (p. 31), the framework of user perceptions that is to be validated in the User Study comprised six constructs. These included perceived comprehensibility, perceived exaggeration, perceived specificity, perceived tentativeness, perceived authenticity, and epistemic beliefs. These conceptual abstractions were converted into equivalent measures to develop the items in the questionnaire (Edwards & Bagozzi, 2000).

Apart from the questions seeking qualitative responses and demographic information, all other questionnaire items sought responses on a five-point rating scale with labels only at extremes anchored by one (strongly disagree) and five (strongly agree). A five-point scale was preferred to dichotomous responses to enhance sensitivity, which refers to the ability of questionnaire items to measure variability in participants' responses (Zikmund, 2003). Besides, scales with labels only at extremes were used because they elicit more accurate responses than scales with labels on all response categories (Weijters et al., 2010). The measures in the questionnaire were evaluated based on parameters such as face validity, content validity, and internal consistency reliability as shown later in Section 5.3.2 (p. 123).

5.3.1. Measures

The questionnaire comprised two segments. The first included items to measure participants' perceptions of each of the three reviews both quantitatively and qualitatively. Given that the Linguistic Study analyzed the linguistic cues of reviews for titles as well as descriptions (cf. Section 3.2, p. 42), the User Study also traced users' perceived linguistic cues for the two components separately. This called for operationalizing the perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness—as formative constructs (Wixom & Watson, 2001). All other constructs were reflective.

Specifically, formative constructs are those that are caused by measures related to differing aspects of the underlying concept. In other words, the measures define the characteristics of the construct. The measures, which are not necessarily correlated with

one another, capture a unique aspect of formative constructs (Diamantopoulos & Winklhofer, 2001; Jarvis et al. 2003; Petter et al., 2007; Wixom & Watson, 2001).

Formative constructs differ from the more common reflective constructs, which are known to affect the measures. In other words, measures are manifestations of reflective construct. Therefore, changes in reflective constructs cause changes in their measures, which are expected to be positively correlated with one another (Diamantopoulos & Winklhofer, 2001; Jarvis et al. 2003; Petter et al., 2007; Wixom & Watson, 2001).

Perceived comprehensibility refers to the extent to which humans consider reviews as being easy to read. It was quantitatively measured by asking participants to rate their degree of agreement with the following statements: “The title of the review is easy to read,” and “The description of the review is easy to read” (Berger & Iyengar, 2013; Galante & Thomson, 2017). Higher scores in these items indicate greater perceived comprehensibility.

Perceived exaggeration refers to the extent to which humans consider reviews to have gone overboard in an attempt to sound convincing. It was quantitatively measured by asking participants to rate their degree of agreement with the following statements: “The title of the review is exaggerated,” and “The description of the review is exaggerated” (Fletcher et al., 2009; Hu et al., 2012). Higher scores in these items indicate greater perceived exaggeration.

Perceived specificity refers to the extent to which humans consider reviews as being informative and detailed. It was quantitatively measured by asking participants to rate their degree of agreement with the following statements: “The title of the review is

informative,” and “The description of the review is informative” (Bond & DePaulo, 2006; Shapiro et al., 1994). Higher scores in these items indicate greater perceived specificity.

Perceived tentativeness refers to the extent to which humans consider a review to be written vaguely with hesitation. It was quantitatively measured by asking participants to rate their degree of agreement with the following statements: “The title of the review appears tentative,” and “The description of the review appears tentative” (Ekman & Friesen, 1969; Latta, 2004). Higher scores in these items indicate greater perceived tentativeness.

Perceived authenticity was quantitatively measured by asking participants to rate their degree of agreement with the following statements: “The review is a genuine account of post-trip experience,” “The review is written after a stay in the hotel,” and “The review is an honest description of a stay in the hotel” (Blair et al., 2010; Kim et al., 2015). Higher scores in these items indicate greater perceived authenticity.

To glean richer insights into participants’ perceptions of reviews and facilitate triangulation (Decrop, 1999), three open-ended questions were used to facilitate a qualitative analysis. Specifically, for every review, participants were asked the following: “Do you think the review is authentic? Why?,” “Did you use review titles to determine review authenticity? If yes, how?,” and “Did you use review descriptions to determine review authenticity? If yes, how?” The items in the first segment of the questionnaire are presented in Table 5.4.

Table 5.4: Items in the first segment of the questionnaire.

Perceptions of Reviews	Questionnaire Items
Perceived comprehensibility (formative construct)	PCom1: The title of the review is easy to read. PCom2: The description of the review is easy to read.
Perceived exaggeration (formative construct)	PExa1: The title of the review is exaggerated. PExa2: The description of the review is exaggerated.
Perceived specificity (formative construct)	PSpe1: The title of the review is informative. PSpe2: The description of the review is informative.
Perceived tentativeness (formative construct)	PTen1: The title of the review appears tentative. PTen2: The description of the review appears tentative.
Perceived authenticity (reflective construct)	PA1: The review is a genuine account of post-trip experience. PA2: The review is written after a stay in the hotel. PA3: The review is an honest description of a stay in the hotel.
Open-ended questions	Quali1: Do you think the review is authentic? Why? Quali2: Did you use review titles to determine review authenticity? If yes, how? Quali3: Did you use review descriptions to determine review authenticity? If yes, how?

The second segment of the questionnaire included items to measure participants' individual differences. The individual difference of epistemic beliefs includes two disparate facets. These include perceived reliability of knowledge, and perceived justification for knowing (Kammerer et al., 2013). The former was measured using 11 items while the latter was measured using four items. Higher scores in these items indicate greater epistemological naivety (Bråten et al., 2005; Hofer, 2004; Hofer & Pintrich, 1997; Kammerer et al., 2013).

Participants' Internet anxiety, familiarity with review websites, and attitude toward reviews for hotels were obtained because these variables served as useful control variables in related works (Kwon & Sung, 2012; McKnight & Kacmar, 2006; Qiu et al., 2012). Internet anxiety was measured by asking participants to rate their degree of agreement with the following statements: "I feel apprehensive about using the Internet," "I hesitate to use the Internet for fear of making mistakes," and "I avoid the Internet because it is

intimidating to me” (McKnight & Kacmar, 2006). Higher scores in these items indicate greater anxiety in using the Internet.

Familiarity with review websites was measured by asking participants to rate their degree of agreement with the following statements: “I am familiar with the use of review websites such as Amazon.com and TripAdvisor.com,” and “I visit review websites such as Amazon.com and TripAdvisor.com” (Blanco et al., 2010). Higher scores in these items indicate greater familiarity with review websites.

Attitude toward reviews for hotels was measured by asking participants to rate their degree of agreement with the following statements: “I generally like reading online user reviews for hotels,” and “I generally find it interesting to read online user reviews for hotels” (Park et al., 2007). Higher scores in these items indicate more positive attitude toward reviews for hotels.

Participants’ demographic information was also sought. Specifically, they were asked to provide four demographics information, namely, gender, age group, educational/employment status, and nationality. The items in the second segment of the questionnaire are presented in Table 5.5.

Table 5.5: Items in the second segment of the questionnaire.

Individual Differences	Questionnaire Items
Epistemic beliefs: Reliability (reflective construct)	EpiR1: Online reviews on the Internet reflect accurate knowledge about hotels. EpiR2: Online reviews provide me with most of the knowledge I need to select a hotel. EpiR3: Correct information about hotels can be found in online reviews. EpiR4: In online reviews, the richness of detail about hotels is most prominent. EpiR5: Online reviews provide correct information about hotels. EpiR6: The most important aspect of online reviews is that they contain specific facts about hotels. EpiR7: I am most confident that I have selected the appropriate hotel when I have used online reviews as the source of information. EpiR8: Most of what is true about hotels are available in online reviews. EpiR9: Online reviews contain concrete information about hotels. EpiR10: The strength of online reviews is the vast amount of detailed information available about hotels. EpiR11: Correct evaluation of hotels is available in online reviews.
Epistemic beliefs: Justification (reflective construct)	EpiJ1: To check the credibility of hotel-related information available in online reviews, I try to compare multiple sources. (R) EpiJ2: I check if the hotel-related information available in online reviews is logical. (R) EpiJ3: To check if the hotel-related information available in online reviews is reliable, I evaluate it in relation to other knowledge I have. (R) EpiJ4: I evaluate hotel-related information available in online reviews by checking more sources. (R)
Internet anxiety (reflective construct)	Anx1: I hesitate to use the Internet for fear of making mistakes. Anx2: I avoid the Internet because it is intimidating to me.
Website familiarity (reflective construct)	Fam1: I am familiar with the use of review websites such as Amazon.com and TripAdvisor.com. Fam2: I visit review websites such as Amazon.com and TripAdvisor.com.
Attitude toward reviews (reflective construct)	Att1: I generally like reading online reviews for hotels. Att2: I generally find it interesting to read online reviews for hotels.
Demographics	D1: Indicate your gender: (a) male, (b) female. D2: Indicate your age group: (a) 21-25 years, (b) 26-35 years, (c) 36-45 years. D3: If you are a student, indicate your current educational status. If you are a working adult, indicate your highest educational qualification: (a) Bachelor's student, (b) Master's student, (c) Doctoral student, (d) Bachelor's, (e) Master's, (f) Ph.D. D4: Indicate your nationality.

Note. (R) indicates reverse-coded items.

5.3.2. *Evaluation of the Measures*

Prior to data collection, the questionnaire items were evaluated based on two parameters, namely, face validity, and content validity. Face validity indicates the degree to which the face value of the questionnaire items shows that they measure what they are

supposed to measure (Tharenou et al., 2007). Informed by prior works (Chen et al., 2014), face validity was examined by pre-testing the questionnaire items using a group of 10 participants. They were graduate students of Information Studies or Information Systems in a large public university, regular readers or contributors of reviews, and had travel experience in the previous year. They were required to comment on the extent to which the questionnaire items represented the measured construct. A high level of consensus suggested face validity.

Content validity refers to the extent to which items correctly represent the theoretical content of the construct. It is generally ensured through in-depth literature review (Casaló et al., 2011; Chen et al., 2014). Therefore, the development of all the questionnaire items was informed by the literature.

Besides, content validity is commonly ascertained using the two-step Q-sorting approach (Petter et al., 2007; Moore & Benbasat, 1991). The first step requires providing a participant, who has no prior knowledge about the study, with sets of all the questionnaire items, and the names of all the involved constructs. The participant is required to match the sets with the appropriate constructs. The second step of Q-sorting requires providing a different participant, who has no prior knowledge about the study, with only the sets of questionnaire items. The participant is required to guess the constructs.

To perform Q-sorting, four participants who had no prior knowledge about the study were recruited. They were graduate students of Information Studies or Information Systems in a large public university, and regular readers of reviews. Two of them were recruited for the first step while the other two were recruited for the second step. Those

recruited for the first step could accurately match the sets of questionnaire items with the constructs. Those recruited for the second step generally offered close synonyms to the actual constructs in the study. For example, the construct “Perceived Comprehensibility” was predicted as “Ease of Reading,” while “Perceived Authenticity” was predicted as “Credibility.” This suggested content validity of the items.

After data collection, the quantitatively measured questionnaire items were further evaluated empirically. Informed by prior works (Bollen & Lennox, 1991; Jarvis et al., 2003; Luo et al., 2012; Petter et al., 2007; Urbach & Ahlemann, 2010), questionnaire items for formative constructs were evaluated for multicollinearity and item weightings while those for reflective constructs were evaluated for internal consistency reliability, convergent validity, and discriminant validity.

For the formative constructs, multicollinearity was assessed based on variance inflation factor while item weightings were obtained through a principal component analysis (Bollen & Lennox, 1991; Luo et al., 2012; Petter et al., 2007; Urbach & Ahlemann, 2010). For the reflective constructs, internal consistency reliability was checked in terms of Cronbach’s alpha and composite reliability (Nunnally, 1978; Tharenou et al., 2007; Werts et al., 1974). Convergent validity was assessed based on average variance extracted (Tharenou et al., 2007; Verhoef et al., 2002). Discriminant validity, also called divergent validity, was checked by examining correlations among the constructs in relation to the square root of average variance extracted (Campbell & Fiske, 1959; Casaló et al., 2011; Fornell & Larcker, 1981).

Finally, all the quantitatively measured questionnaire items were evaluated based on the parameter of common method bias. Common method bias is a condition that occurs

when a single factor explains bulk of the variance in a research framework. It is detected using Herman's single factor test (Podsakoff & Organ, 1986). Table 5.6 summarizes the methods employed to evaluate the measures.

Table 5.6: Methods employed for evaluation of measures.

Time of Evaluation	Evaluation Parameters	Methods of Evaluation
Before collecting data	Face validity	Questionnaire pre-testing (Chen et al., 2014)
	Content validity	In-depth literature review followed by two-step Q-sorting (Casaló et al., 2011; Moore & Benbasat, 1991)
After collecting data	Multicollinearity ¹	Variance inflation factor (Bollen & Lennox, 1991; Luo et al., 2012)
	Item weightings ¹	Principal component analysis (Petter et al., 2007; Urbach & Ahlemann, 2010)
	Internal consistency reliability ²	Cronbach's alpha and composite reliability (Nunnally, 1978; Werts et al., 1974)
	Convergent validity ²	Average variance extracted (Verhoef et al., 2002)
	Discriminant validity ²	Inter-construct correlation in relation to the square root of average variance extracted (Casaló et al., 2011; Fornell & Larcker, 1981)
	Common method bias	Herman's single factor test (Podsakoff & Organ, 1986)

Note. ¹ indicates parameters evaluated only for formative constructs.

² indicates parameters evaluated only for reflective constructs.

5.4. Analytical Approach

5.4.1. Quantitative Analysis

The quantitative data were analyzed using structural equation modeling (SEM). It was preferred to multiple regression analysis because SEM is capable of analyzing latent

variables (constructs) measured using multiple observed variables (questionnaire items) while recognizing measurement errors (Gefen et al., 2011; Zhang et al., 2014). In contrast, multiple regression analysis only deals with observed variables. As shown earlier in Section 2.3.2 (p. 31), this study sought to examine relations among latent variables (e.g., relation between constructs such as perceived comprehensibility and perceived authenticity), and not those that are directly observable.

There are two broad approaches to SEM, namely, covariance-based SEM, and variance-based SEM. Covariance-based SEM refers to approaches that use maximum likelihood estimation to assess the fit between a theoretical model and a covariance matrix (Ranganathan & Sethi, 2002; Ridgon, 1998). It is available in statistical packages such as LISREL and EQS.

In contrast, variance-based SEM refers to approaches that use partial least squares estimation to minimize residual variances of the variables (Chin, 1998; Vilares et al., 2010; Wold & Jöreskog, 1982; Wold, 1985). It is available in statistical packages such as SmartPLS and PLS Graph.

Although the use of covariance-based SEM is complementary to that of variance-based SEM, the latter was more appropriate for this study due to three reasons. First, covariance-based SEM is suited to validate existing theories while variance-based SEM is suited for theory development and prediction (Fornell & Bookstein, 1982; Gefen et al., 2000; Wold, 1985). This study does not validate a single existing theory. Rather, it seeks to test a framework developed based on the literature on deception and information processing as explained earlier in Section 2.3 (p. 23).

Second, for covariance-based SEM, variables need to follow a multivariate normal distribution, the sample has to be large, and there has to be independence of observations. In contrast, variance-based SEM is a distribution free approach that does not require such assumptions (Bollen, 1989; Urbach & Ahlemann, 2010). Given that this study collects data to test a newly built framework, the stringent requirements of covariance-based SEM are not guaranteed to be met (Ranganathan & Sethi, 2002; Vilares et al., 2010).

Third, compared with covariance-based SEM, variance-based SEM is better suited to deal with formative constructs (Choudhury & Karahanna, 2008) as well as moderators (Limayem et al., 2007). The framework of user perceptions included not only formative constructs (e.g., perceived comprehensibility) but also a moderator (e.g., epistemic beliefs).

For variance-based SEM, the minimum sample size equals 10 times the greater of the number of items comprising the most formative construct, or the number of independent constructs influencing a single endogenous dependent construct (Barclay et al., 1995; Chin, 1998; Wixom & Watson, 2001). In the framework of user perceptions, the number of items measuring each of the formative constructs is two as shown earlier in Table 5.4 (p. 121). Therefore, the most formative construct was measured by two items. Moreover, as evident from Figure 2.2 (p. 35), the number of independent constructs influencing a single dependent construct—perceived authenticity—is four. Therefore, the minimal sample size equals 40 (10 times the greater of two or four). For the User Study, a sample size of at least 300 was targeted. Eventually, complete responses from 380 participants were obtained. This exceeds the recommended threshold.

To address the research objective RO 2.1 (indicated earlier in Figure 1.1, p. 9), the coefficients of the paths from the independent constructs—perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness—to the dependent construct perceived authenticity in the SEM model were examined in terms of t-statistics, statistical significance, magnitude and direction. The goal was to identify the perceived linguistic cues that were significant in predicting perceived authenticity. To check for statistical significance, the bootstrap procedure with 500 resamples was employed (Chen et al., 2014; Ringle et al., 2005; Wixom & Watson, 2001).

Thereafter, the model was tested for its explanatory power, magnitude or importance of statistical relations, and goodness-of-fit. Specifically, the variability in the endogenous construct perceived authenticity that is explained by the perceived linguistic cues was estimated in terms of the coefficient of determination, denoted as R^2 (Chin, 1998; Urbach & Ahlemann, 2010). The magnitude or the importance of the identified statistical relations was assessed based on effect size, denoted by f^2 (Cohen, J., 1988). Finally, the goodness-of-fit was ascertained using Stone-Geisser's blindfolding procedure in terms of Q^2 (Tenenhaus et al., 2004).

To address the research objective RO 2.2 (indicated earlier in Figure 1.1, p. 9), the model was tested with epistemic beliefs as a moderator. To test for moderation, a multi-group analysis is commonly conducted (Casaló et al., 2011; Velayutham et al., 2012). Data are split into two groups—high and low—based on levels of the moderating variable. Next, the analysis is run separately on each group of data. Thereafter, the group comparison method is employed to determine whether the differences in path coefficients

between the two groups are statistically significant (Chin, 2000; Keil et al., 2000; Zhang et al., 2014).

As highlighted earlier in Section 5.3.1 (p. 121), epistemic beliefs were operationalized as two disparate facets, namely, perceived reliability of knowledge, and perceived justification for knowing (Kammerer et al., 2013). The questionnaire items measuring perceived reliability of knowledge as well as those measuring perceived justification for knowing were used separately to categorize the participants as either epistemologically naïve or epistemologically robust.

With respect to perceived reliability of knowledge, the participants were categorized as either naïve or robust according to the arithmetic mean of their aggregated responses to the questionnaire items measuring the construct (EpiR1 to EpiR11 shown in Table 5.5, p. 123). With respect to perceived justification for knowing, the participants were categorized as either naïve or robust according to the arithmetic mean of their aggregated responses to the questionnaire items measuring the construct (EpiJ1 to EpiJ4 shown in Table 5.5, p. 123). Individuals within $\pm \frac{1}{2}$ standard deviation around the mean were eliminated to confine the scope of the multi-group analysis to participants who were clearly epistemologically naïve or robust in terms of perceived reliability of knowledge (Casaló et al., 2011; Rodríguez et al., 2008).

The PLS analysis was conducted separately four times for (1) the participants who were epistemologically naïve with respect to perceived reliability of knowledge, (2) for those who were epistemologically robust with respect to perceived reliability of knowledge, (3) for those who were epistemologically naïve with respect to perceived justification for knowing, and (4) for those who were epistemologically robust with

respect to perceived justification for knowing. Next, the group comparison method was employed to identify differences between epistemologically naïve and robust participants with respect to perceived reliability of knowledge. Another round of the group comparison method was used to identify differences between epistemologically naïve and robust participants with respect to perceived justification for knowing (Chin, 2000; Keil et al., 2000; Zhang et al., 2014).

All variance-based SEM analyses were conducted using SmartPLS (Ringle et al., 2005). Participants' age group, gender, educational/employment status, Internet anxiety, familiarity with review websites, and attitude toward reviews as well as the website versions to which individuals were assigned were controlled in the analyses.

5.4.2. Qualitative Analysis

To offer richer insights into the results of RO 2.1 and RO 2.2, qualitative content analysis was conducted on the open-ended responses submitted by the participants. Compared with quantitative analysis, this analytical approach was more exploratory to allow gleaning richer insights into participants' perceptions of review authenticity (Creswell, 2012). Employing this approach was necessary because integrating quantitative and qualitative data to examine research issues at stake is increasingly encouraged. Such a methodological integration offers insights that could not otherwise be obtained (Bryman, 2007; Molina-Azorin, 2016).

The unit of analysis was the open-ended response provided by each of the 380 participants about each of the three reviews that they had seen during participation. Hence, there could be a maximum of 1,140 units of analysis (380 participants x 3

reviews). Of these, there were 22 nonsensical responses, which were eliminated. The remaining 1,118 responses (1140 - 22) were admitted for an inductive content analysis.

For this purpose, three independent coders were recruited. They were graduate students of Information Systems in a large public university, regular readers or contributors of reviews, and had travel experience in the previous year. The inductive content analysis was conducted for a period of almost six months. To minimize fatigue-induced coding errors, it was limited to at most one hour of coding per session as well as to at most three sessions per day.

The analysis procedure involved three steps. First, the coders conducted a line-by-line scrutiny of the open-ended responses to generate themes inductively. Specifically, the responses were first microanalyzed to identify emergent themes based on how participants voiced their perceptions of reviews. Due to the interpretative nature of such an analysis, the researcher allowed the coders to annotate a single statement from participants into multiple codes where necessary (Glaser & Strauss, 1967; Strauss & Corbin, 1998).

Second, the coders communicated the preliminary findings to the researcher. The coders along with the researcher examined the open-ended responses in face-to-face meetings. The emergent themes were constantly compared against one another to ascertain inter-relationships (Glaser & Strauss, 1967; Strauss & Corbin, 1998). When needed, similar themes were collapsed. Development and enhancement of themes continued until new observations failed to add significantly to existing themes, thereby suggesting theoretical saturation. Such an approach of analysis is informed by the literature (Papathanassis & Knolle, 2011; Sanderson, 2013). To confirm the nature of the

existing themes, a coding scheme was developed (Saldaña, 2015), which is provided later in Table 6.8 (p. 162).

Third, the coders used the coding scheme to revisit the open-ended responses using a more confirmatory approach compared with the previous steps. The purpose was not only to confirm the identified themes but also to look for any other themes that might have been overlooked. Such revisits are known to enhance consistency of content analysis (Papathanassis & Knolle, 2011; Fileri, 2016). This was especially important given the explorative nature of the analysis that prevented the use of an a priori codebook and an inter-coder reliability analysis (Burla et al., 2008; Papathanassis & Knolle, 2011; Sanderson, 2013). Nonetheless, this approach is inspired by the recommendation that “the most significant and lasting contributions have been made by researchers who employed an often loose qualitative methodology” (Cohen, E., 1988, p. 30). The overall analytical procedure is summarized in Figure 5.4.

Figure 5.4: Summary of the analytical procedure for the User Study.

5.5. Methods of the Robustness Check

The Robustness Check, which sought to confirm the findings of this research, investigated the extent to which humans could be taught how to discern review authenticity. Given that guidance on ways to evaluate information improves individuals' information-processing strategies (Argelagós & Pifarré, 2012; Munzel, 2015), the Robustness Check examined the extent to which a guideline could help humans identify linguistic differences between authentic and fictitious reviews. It included three tasks: development of the guideline, pre-test of the guideline, and efficacy test of the guideline.

For the first task of developing the guideline, the results of the Linguistic Study reported in Section 4.3 (p. 81) were leveraged. Specifically, the results demonstrated that exaggeration had the highest number of variables (19 variables) that helped distinguish between authentic and fictitious reviews, followed by specificity (9 variables). However, comprehensibility and tentativeness had fewer such variables (5 variables each). This indicates that exaggeration offered maximal scope to identify fictitious reviews, followed by specificity. On the other hand, comprehensibility and specificity offered relatively less opportunity to distinguish between authentic and fictitious reviews. Hence, in order to discern review authenticity, the linguistic cues could be leveraged in the following order—exaggeration, specificity, followed by comprehensibility or tentativeness.

Maintaining this order, the differences between authentic and fictitious reviews that were consistently prominent across both titles and descriptions are listed in Table 5.7. Specifically, there are six exaggeration-related variables, three specificity-related variables, and one comprehensibility-related and one tentativeness-related variable. These

are derived from the set of all differences between authentic and fictitious reviews listed earlier in Table 4.12 (p. 103).

Only the differences listed in Table 5.7 between authentic and fictitious reviews that were evident consistently across titles as well as descriptions were utilized to develop the guideline for the Robustness Check. This obviated the need to include separate sets of instructions in the guideline pertaining to titles and descriptions of reviews. Such instructions would have hindered ease of understanding.

Table 5.7: Differences consistently prominent across review titles and descriptions.

Linguistic Cues	Variables	Authentic Reviews	Fictitious Reviews
Exaggeration	Negative emotion words	Scanty	Abundant
	Emotiveness	Low	High
	Firm words	Scanty	Abundant
	Hotel names	Scanty	Abundant
	Exclamation marks	Scanty	Abundant
	Function words	Scanty	Abundant
Specificity	Nouns	Abundant	Scanty
	Pronouns	Scanty	Abundant
	Spatial words	Abundant	Scanty
Comprehensibility	Long words	Scanty	Abundant
Tentativeness	First person singular words	Scanty	Abundant

The guideline was developed to resemble a decision-tree with three decision-points as shown in Figure 5.5. Specifically, a decision-tree was chosen over a linear list of cues because the former was unanimously found to be cognitively more manageable by 10 participants, who were recruited for a pilot study based on the same eligibility criteria reported earlier in Section 5.2 (p. 112). Their feedback suggested that a decision-tree could be more efficacious to discern review authenticity than a linear list of cues.

Figure 5.5: Decision-tree format of the guideline to discern review authenticity.

At the first decision-point of the decision-tree, the guideline required individuals to rely on exaggeration to identify fictitious reviews. If exaggeration cues were unavailable, it required individuals to check reviews' specificity at the second decision-point to spot fictitious entries. If specificity cues were unavailable, the guideline required individuals to examine reviews' comprehensibility or tentativeness to find fictitious reviews. Authentic reviews were left to be labelled by elimination. Put differently, the guideline prioritized accurate identification of fictitious reviews over that of authentic entries. This was necessary to minimize the chances of labelling fictitious reviews as authentic. After all, consequences are direr when users regard a fictitious review authentic than an authentic review fictitious (Chen & Lin, 2013).

Each decision-point in the guideline was presented as an instruction. The instruction for the first decision-point required individuals to check if a review was rich in

emotions, especially negative emotion words such as “bad,” firm words such as “never,” hotel names, function words such as “are,” or exclamation marks. No instruction corresponding to the use of hotel names was included because, as indicated earlier in Section 5.1.2 (p. 110), reviews containing hotel names were avoided in the 54 reviews used in this study. If exaggeration cues were available in a review, the guideline required individuals to annotate it as fictitious. Otherwise, they could proceed to the next decision-point.

The instruction for the second decision-point required individuals to check if the review failed to provide details through nouns such as “room” or spatial words such as “location,” and was vague by describing personal experiences using pronouns such as “you.” If yes, it should be annotated as fictitious. Otherwise, they could proceed to the next decision-point.

The instruction for the third decision-point required individuals to check if the review used long words such as “claustrophobic,” or if it was rich in first person singular words such as “me.” If yes, it should be annotated as fictitious. Otherwise, it could be labelled as authentic by elimination.

The second task in the Robustness Check involved pre-testing the guideline. For this purpose, its instructions were examined for clarity using separate batches of 10 participants who were recruited on meeting the eligibility criteria reported earlier in Section 5.2 (p. 112). In one-to-one meetings with the researcher, the participants were required to think aloud while going through the instructions. They were asked to comment on its ease of understanding.

In an initial round of pre-test, the instructions for the decision-points were corroborated with several excerpts of authentic and fictitious reviews. The participants however unanimously complained about information overload. Based on their feedback, excerpts were completely removed. To make a trade-off between managing participants' cognitive load and ensuring the efficacy of the guideline, instructions for the decision-points were revised to highlight only selected word samples.

In the next round of pre-test, the comments of the participants were more favorable. However, two participants complained about confusions arising from inconsistencies in the instructions. They pointed that while some instructions were of the form, "Check if the review is rich in . . . If yes, annotate it as fake." other instructions stated, "Check if the review lacks . . . If yes, annotate it as fake." Therefore, the instructions were fine-tuned to maintain a consistent tone with sentences of the form, "Check if the review is rich in . . . If yes, annotate it as fake." In this way, the instructions were now more understandable by consistently asking participants to look for confirming evidences rather than a mixture of both confirming and disconfirming evidences.

However, this change meant that the linguistic differences in terms of nouns and spatial words could not be included in the guideline. Both were more abundant in authentic reviews than fictitious ones. Highlighting these differences would have given rise to instructions asking annotators to look for disconfirming evidence as in, "Check if the review lacks nouns and spatial words. If yes, annotate it as fake." In any case, reviews rich in nouns and spatial words could still be identified as authentic by elimination.

In another round of pre-test, all the participants were able to follow the fine-tuned instructions without any ambiguity. The guideline with these instructions was finalized. It is provided in Appendix G (p. 207).

The third task in the Robustness Check involved testing the efficacy of the guideline. For this purpose, the guideline was used as an intervention in a between-participants experimental design. Experiments are appropriate to investigate the effect of some independent factors on an outcome under controlled settings (Highhouse, 2009). In this case, the independent factor was the intervention while perceived review authenticity was the outcome. In particular, 240 annotators were recruited for the experiment based on the eligibility criteria specified earlier in Section 5.2 (p. 112). They were randomly assigned to one of the two conditions: without intervention (henceforth, control group), or with intervention (henceforth, treatment group). They were required to predict the authenticity of each of the 54 reviews (cf. Section 5.1.2, p. 110) that were arranged in a random sequence.

Efforts were made to have several annotators labelling a manageable volume of reviews. Related prior works often required annotators to label in excess of 100 reviews each (Lau et al., 2011; Li et al., 2011). In contrast, the Robustness Check required annotators to label 54 reviews each. This helped lower annotators' cognitive load, thereby minimizing chances of fatigue-induced errors. Moreover, unlike the previous works (Lau et al., 2011; Li et al., 2011), this research allowed the annotation to be conducted systematically in a true experimental design that included post-test measures, control group, and random assignment to control or treatment groups.

Annotators in the control group were asked to heuristically determine if reviews were authentic or fictitious to the best of their abilities. Those in the treatment group were asked to follow the instructions in the intervention to discern review authenticity. Taking the cue from prior works (Du et al., 2013), the annotators were kept unapprised of the exact numbers of authentic and fictitious reviews. This ensured that they could not reverse-engineer to complete the annotation task. All annotators received SGD 5 as a token of appreciation. Finally, the extent to which the annotators in the control group differed from those in the treatment group in accurately discerning review authenticity was examined using independent samples t-tests. The tasks in the Robustness Check are depicted in Figure 5.6.

Figure 5.6: Tasks in the Robustness Check.

5.6. Summary of the User Study - Methods

This chapter has elaborated the methods that were used to conduct the User Study. Specifically, it began by describing the stimuli that was used in the survey. Next, it explained the survey procedure, development of the questionnaire, and the approaches for

data analysis. Thereafter, it presented the methods employed to conduct the Robustness Check.

The survey stimuli included a simulated review website LoveToTravel.com. A set of 54 reviews (27 authentic + 27 fictitious) was identified to be displayed on the website. Participants' perceptions of reviews as well as information about their individual differences were captured using a questionnaire. Analytical approaches involved variance-based SEM and content analysis. For the robustness check, a guideline was developed to help humans distinguish between authentic and fictitious reviews. After a few rounds of pre-test, the efficacy of the guideline was checked using an experimental design.

Having explained the methods of data collection and analyses for the User Study, this chapter has laid the foundation to present its findings, which is the purpose of the next chapter. It describes the findings of the two research objectives—RO 2.1 and RO 2.2—corresponding to RQ 2 followed by those of the Robustness Check.

CHAPTER SIX USER STUDY – FINDINGS

This chapter presents the findings of the User Study that sought to answer the research question RQ 2: To what extent are users able to perceive the authenticity of reviews? The investigation was underpinned by the framework of user perceptions that drew from the theoretical perspective of information processing. It entailed addressing the two research objectives—RO 2.1 and RO 2.2—followed by the Robustness Check.

The first research objective (RO 2.1) attempted to analyze the role of humans' perceived linguistic cues in explaining perceived authenticity of reviews. To address this objective, the data collected from the online survey described earlier in Section 5.2 (p. 112) were analyzed using partial least squares structural equation modeling (PLS). Perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness were the independent constructs while perceived authenticity was the dependent construct.

The second research objective (RO 2.2) attempted to examine the extent to which epistemic beliefs moderated the relation between perceived linguistic cues and perceived authenticity of reviews. To address this objective, a multi-group PLS analysis was conducted. One group included responses from epistemologically naïve participants while another group comprised responses from epistemologically robust participants. To glean richer insights into RO 2.1 and RO 2.2, qualitative analysis was conducted on the participants' responses to the open-ended questions in the questionnaire (cf. Table 5.4, p. 121).

Finally, the Robustness Check was conducted to confirm the findings of this research. It investigated the extent to which humans could be taught how to discern review authenticity. As indicated earlier in Section 5.5 (p. 134), a guideline was developed to help individuals identify linguistic differences between authentic and fictitious reviews. The efficacy of the guideline was examined with the help of human annotators in an experimental design.

As shown in Figure 6.1, the remainder of the chapter is organized into seven sections. The first section (p. 144) summarizes the descriptive statistics of the dataset used in this study. The second section (p. 151) presents the results pertaining to the role of perceived linguistic cues in explaining perceived authenticity of reviews in order to address RO 2.1. The third section (p. 154) presents the results pertaining to the moderating role of epistemic beliefs in order to address RO 2.2. Specifically, the moderating role is examined for perceived reliability of knowledge (p. 154), and perceived justification for knowing (p. 157), which are two disparate facets of epistemic beliefs (Kammerer et al., 2013). The fourth section (p. 161) describes the results of the qualitative analysis. The fifth section (p. 170) presents the results of the Robustness Check. The sixth section (p. 172) discusses the results in relation to the literature. Finally, the seventh section (p. 176) summarizes the chapter.

Figure 6.1: Overview of Chapter Six.

6.1. Descriptive Statistics of the User Dataset

As shown earlier in Table 5.3 (p. 117), data for the User Study included responses from 380 participants. Of these, 214 were males while the remaining 166 were females. Majority of the participants aged between 21 and 25 years, and were undergraduate students. The participants straddled across 16 nationalities. The top three most represented nationalities included Singaporean, Indian, and mainland Chinese. Some of the other nationalities included Americans, Ecuadorians, Italians and Sri Lankans. The demographic details of the participants are summarized in Table 6.1.

Table 6.1: Demographic details of the participants.

	Frequency	Percentage	Cumulative
Gender			
Male	214	56.31%	214 (56.31%)
Female	166	43.69%	380 (100%)
Age group			
21-25 years	225	59.21%	225 (59.21%)
26-35 years	137	36.05%	362 (95.26%)
36-45 years	18	4.74%	380 (100%)
Educational/Employment status			
Student pursuing Bachelor's degree	143	37.63%	143 (37.63%)
Student pursuing Master's degree	63	16.58%	206 (54.21%)
Student pursuing Doctoral degree	37	9.74%	243 (63.95%)
Working adult with Bachelor's degree	56	14.74%	299 (78.69%)
Working adult with Master's degree	65	17.10%	364 (95.79%)
Working adult with Doctoral degree	16	4.21%	380 (100%)
Nationality			
Singaporean	150	39.47%	150 (39.47%)
Indian	96	25.26%	246 (64.73%)
Chinese	54	14.21%	300 (78.94%)
Others	80	21.05%	380 (100%)

Note. N = 380.

As indicated earlier in Section 5.3.1 (p. 118), the questionnaire for the User comprised two segments. The first included items to measure the participants' perceptions of reviews. Descriptive statistics of the quantitatively measured items in this segment of the questionnaire—perceived comprehensibility (PCom), perceived exaggeration (PExa), perceived specificity (PSpe), perceived tentativeness (PTen), and perceived authenticity (PA)—are presented in Table 6.2.

Table 6.2: Descriptive statistics of items in the first segment of the questionnaire.

Perceptions of Reviews	Questionnaire Items	Mean	SD
Perceived comprehensibility (formative construct)	PCom1: The title of the review is easy to read.	4.14	0.96
	PCom2: The description of the review is easy to read.	3.92	0.93
Perceived exaggeration (formative construct)	PExa1: The title of the review is exaggerated.	2.94	1.18
	PExa2: The description of the review is exaggerated.	3.01	1.19
Perceived specificity (formative construct)	PSpe1: The title of the review is informative.	3.46	1.16
	PSpe2: The description of the review is informative.	3.64	1.05
Perceived tentativeness (formative construct)	PTen1: The title of the review appears tentative.	2.98	1.02
	PTen2: The description of the review appears tentative.	2.96	0.99
Perceived authenticity (reflective construct)	PA1: The review is a genuine account of post-trip experience.	3.33	1.20
	PA2: The review is written after a stay in the hotel.	3.44	1.17
	PA3: The review is an honest description of a stay in the hotel.	3.32	1.16

The second segment of the questionnaire included items to measure the participants' individual differences. Descriptive statistics of the quantitatively measured items in this segment of the questionnaire—epistemic beliefs with respect to perceived reliability of knowledge (EpiR), epistemic beliefs with respect to perceived justification for knowing (EpiJ), Internet anxiety, website familiarity, and attitude toward reviews—are presented in Table 6.3.

Table 6.3: Descriptive statistics of items in the second segment of the questionnaire.

Individual Differences	Questionnaire Items	Mean	SD
Epistemic beliefs: Reliability (reflective construct)	EpiR1: Online reviews on the Internet reflect accurate knowledge about hotels.	3.28	0.87
	EpiR2: Online reviews provide me with most of the knowledge I need to select a hotel.	3.74	0.85
	EpiR3: Correct information about hotels can be found in online reviews.	3.78	0.80
	EpiR4: In online reviews, the richness of detail about hotels is most prominent.	3.72	0.94
	EpiR5: Online reviews provide correct information about hotels.	3.48	0.83
	EpiR6: The most important aspect of online reviews is that they contain specific facts about hotels.	3.89	0.82
	EpiR7: I am most confident that I have selected the appropriate hotel when I have used online reviews as the source of information.	3.42	0.90
	EpiR8: Most of what is true about hotels are available in online reviews.	3.54	0.83
	EpiR9: Online reviews contain concrete information about hotels.	3.40	0.88
	EpiR10: The strength of online reviews is the vast amount of detailed information available about hotels.	3.74	0.91
	EpiR11: Correct evaluation of hotels is available in online reviews.	3.70	0.82
Epistemic beliefs: Justification (reflective construct)	EpiJ1: To check the credibility of hotel-related information available in online reviews, I try to compare multiple sources. (R)	1.67	0.78
	EpiJ2: I check if the hotel-related information available in online reviews is logical. (R)	1.84	0.79
	EpiJ3: To check if the hotel-related information available in online reviews is reliable, I evaluate it in relation to other knowledge I have. (R)	1.88	0.78
	EpiJ4: I evaluate hotel-related information available in online reviews by checking more sources. (R)	1.84	0.80
Internet anxiety (reflective construct)	Anx1: I hesitate to use the Internet for fear of making mistakes.	2.09	1.11
	Anx2: I avoid the Internet because it is intimidating to me.	1.73	0.94
Website familiarity (reflective construct)	Fam1: I am familiar with the use of review websites such as Amazon.com and TripAdvisor.com.	4.04	0.91
	Fam2: I visit review websites such as Amazon.com and TripAdvisor.com.	3.76	1.02
Attitude toward reviews (reflective construct)	Att1: I generally like reading online reviews for hotels.	3.78	0.93
	Att2: I generally find it interesting to read online reviews for hotels.	3.76	0.92

Note. (R) indicates reverse-coded items. Means and SDs are reported after reversal. In other words, higher values indicate greater epistemological naivety and vice-versa.

As indicated earlier in Section 5.3.1 (p. 118), there were four formative constructs.

These included perceived comprehensibility (PCom), perceived exaggeration (PExa),

perceived specificity (PSpe), and perceived tentativeness (PTen) of reviews. Formative

constructs are those that are caused by their measures (Diamantopoulos & Winklhofer, 2001; Jarvis et al. 2003; Petter et al., 2007; Wixom & Watson, 2001).

Additionally, there were six reflective constructs. These included perceived authenticity (PA), epistemic beliefs with respect to perceived reliability of knowledge (EpiR), epistemic beliefs with respect to perceived justification for knowing (EpiJ), Internet anxiety (Anx), website familiarity (Fam), and attitude toward reviews (Att). Reflective constructs are those that affect their measures (Diamantopoulos & Winklhofer, 2001; Jarvis et al. 2003; Petter et al., 2007; Wixom & Watson, 2001).

The formative constructs were evaluated in terms of multicollinearity as well as item weightings obtained through a principal component analysis (Bollen & Lennox, 1991; Luo et al., 2012; Petter et al., 2007; Urbach & Ahlemann, 2010). Table 6.4 summarizes the result of this evaluation. Multicollinearity was checked in terms of variance inflation factor (VIF). All values of VIF were less than the recommended threshold 3.3, suggesting that the formative constructs were not impaired by multicollinearity (Diamantopoulos & Siguaw, 2006).

Item weightings are commonly assessed based on statistical significance or coefficients. Item weightings that are statistically significant at the $p < 0.05$ level (Petter et al., 2007), or those with coefficients greater than 0.10 (Lohmöller, 1989; Urbach & Ahlemann, 2010) are deemed important. All measures met the recommended criterion.

Additionally, the reflective constructs were evaluated for internal consistency reliability, convergent validity, and discriminant validity (Casaló et al., 2011; Fornell & Larcker, 1981; Tharenou et al., 2007). Table 6.5 summarizes the result of this evaluation. For all the reflective constructs, values of Cronbach's alpha (α) exceeded 0.70 (Nunnally,

1978), and those of composite reliability (CR) exceeded 0.65 (Steenkamp & Geyskens, 2006; Smith, 1974). Thus, the constructs demonstrated internal consistency reliability.

Convergent validity was checked in terms of average variance extracted (AVE). For all the constructs, values of AVE exceeded 0.40, suggesting sufficient convergent validity (Verhoef et al., 2002).

Discriminant validity was checked by examining the correlations among the constructs in relation to the square roots of AVE. Correlations not equal to one suggest discriminant validity (Casaló et al., 2011). Moreover, square roots of AVE (shaded in Table 6.5) for all the constructs exceeded their correlations with other constructs, further confirming discriminant validity (Fornell & Larcker, 1981).

Table 6.4: Evaluation of formative constructs.

Constructs	Measures	VIF	Weightings
Perceived comprehensibility	PCom1	1.51	0.14
	PCom2	1.60	0.92*
Perceived exaggeration	PExa1	1.53	0.19
	PExa2	1.69	0.88*
Perceived specificity	PSpe1	1.55	0.12
	PSpe2	1.75	0.93*
Perceived tentativeness	PTen1	1.53	0.21
	PTen2	1.72	1.14*

Note. * represents statistically significant item weighting at the $p < 0.05$ level.

Table 6.5: Evaluation of reflective constructs.

Constructs	α	CR	AVE	(1)	(2)	(3)	(4)	(5)	(6)
Perceived authenticity (1)	0.93	0.95	0.87	0.93					
Epistemic beliefs: Reliability (2)	0.89	0.91	0.48	0.15	0.69				
Epistemic beliefs: Justification (3)	0.83	0.89	0.67	-0.08	-0.21	0.82			
Internet anxiety (4)	0.85	0.93	0.87	-0.05	-0.19	0.35	0.93		
Website familiarity (5)	0.80	0.90	0.82	0.03	0.30	-0.31	-0.27	0.90	
Attitude toward reviews (6)	0.87	0.94	0.88	0.07	0.40	-0.22	-0.16	0.52	0.94

Note. Shaded cells at diagonals of the correlation matrix part of the table indicate square roots of AVE.

Next, common method bias among all the quantitatively measured questionnaire items was checked empirically using Herman's single factor test (Podsakoff & Organ, 1986). It requires conducting an exploratory factor analysis with all items measuring the constructs in a research framework. The principal component with unrotated factor solution extracted multiple factors as suggested by the scree plot as well as eigenvalues greater than one. Given that a single factor failed to explain bulk of the variance, common method bias was not a concern (Podsakoff et al., 2003). Additional details of the dataset such as inter-construct correlations, item weightings as well as item loadings coupled with their statistical significance, and cross-loadings are available in Appendix H (p. 209). Table 6.6 summarizes the results of evaluating the measures.

Table 6.6: Results of evaluating the measures.

Evaluation Parameters	Results of Evaluation
Multicollinearity ¹	VIF < 3.3 (Bollen & Lennox, 1991; Luo et al., 2012)
Item weightings ¹	Item weightings statistically significant at the $p < 0.05$ level (Petter et al., 2007), or with coefficients greater than 0.10 (Lohmöller, 1989; Urbach & Ahlemann, 2010)
Internal consistency reliability ²	$\alpha > 0.70$ (Nunnally, 1978), and CR > 0.65 (Smith, 1974)
Convergent validity ²	AVE > 0.40 (Verhoef et al., 2002)
Discriminant validity ²	$\sqrt{\text{AVE}} > \text{inter-construct correlations}$ (Fornell & Larcker, 1981)
Common method bias	Herman's single factor test extracted multiple factors (Podsakoff & Organ, 1986)

Note. ¹ indicates parameters evaluated only for formative constructs.

² indicates parameters evaluated only for reflective constructs.

As explained earlier in Section 5.2 (p. 112), each of the 380 participants were shown three reviews. This meant that reviews were evaluated 1,140 times ($380 \times 3 = 1140$). Across these evaluations, the participants were 68.44% accurate in identifying

authentic reviews, and 43.06% accurate in identifying fictitious entries. The overall accuracy was 55.61%. This is consistent with the nearly axiomatic finding in the literature that human ability to detect deception is not too promising. Their performance could at best be marginally above chance (Bond & DePaulo, 2006; Porter & ten Brinke, 2008).

6.2. Role of Perceived Linguistic Cues

To examine the extent to which perceived linguistic cues could predict perceived authenticity of reviews, PLS analysis was conducted on the data collected from all the 380 participants. Based on the framework of user perceptions, perceived authenticity of reviews was the dependent variable. The independent variables included perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness. Age group, gender, educational/employment status, Internet anxiety, familiarity with review websites, and attitude toward reviews as well as the website versions to which participants were assigned were controlled. None emerged significant.

Figure 6.2 demonstrates the role of perceived linguistic cues in explaining participants' perceived authenticity of reviews. Specifically, perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.39$, $t = 3.69$, $p < 0.001$), thereby lending support to the proposition P6 (cf. Table 2.4, p. 36). This result indicates that reviews that were exaggerated were not perceived as being authentic.

Additionally, perceived specificity was positively related to perceived authenticity ($\beta = 0.34$, $t = 3.00$, $p < 0.01$), thereby lending support to the proposition P7 (cf. Table 2.4, p. 36). This result indicates that reviews that were specific were perceived as being authentic.

However, perceived comprehensibility and perceived tentativeness of reviews were not significantly related to perceived authenticity. Stated otherwise, the comprehensibility and the tentativeness of reviews did not inform participants' perceptions of authenticity. Hence, propositions P5 and P8 (cf. Table 2.4, p. 36) were not supported.

Note. *** represents statistically significant relation at the $p < 0.001$ level.

** represents statistically significant relation at the $p < 0.01$ level.

Sample size = 380.

Number of data points = 1,140 (380 x 3) because each participant evaluated three reviews.

Figure 6.2: PLS result showing the role of perceived linguistic cues.

The coefficient of determination (R^2) for perceived authenticity was 37.50%. In other words, the four perceived linguistic cues explained 37.50% variability in participants' perceptions of review authenticity. In general, R^2 values above 67%, 33%

and 19% are considered substantial, adequate and weak respectively (Chin, 1998; Urbach & Ahlemann, 2010). This suggests that the perceived linguistic cues could adequately explain perceived authenticity.

Nonetheless, the non-substantial R^2 value could be because participants were not adept in relying on the linguistic cues to discern review authenticity. This also offers a possible reason why human ability to discern information authenticity has been consistently found as being far from promising (Bond & DePaulo, 2006; Vrij & Baxter, 1999). Perhaps, humans need to be guided to better distinguish between authentic and fictitious information.

The performance of the model was further examined in terms of effect size (f^2). Specifically, the effect size was 0.60. Values of f^2 between 0.02 and 0.15, between 0.15 and 0.35, and above 0.35 indicate small, medium, and large effect sizes respectively (Chin 1998; Cohen, J., 1988; Gefen et al., 2000). This suggests that the overall magnitude of the statistically significant relations was substantial in explaining perceived authenticity of reviews.

Finally, Stone-Geisser's blindfolding procedure was employed to evaluate goodness-of-fit in terms of Q^2 (Tenenhaus et al., 2004). A positive Q^2 ($Q^2 > 0$) for the endogenous construct perceived authenticity with an omission distance of seven confirmed the predictive relevance of the PLS model (Urbach & Ahlemann, 2010). Stated otherwise, the PLS model was found to fit reasonably well with the data.

6.3. Moderating Role of Epistemic Beliefs

The framework of user perceptions called for examining the extent to which individuals' epistemic beliefs could moderate the relation between perceived linguistic cues and perceived authenticity. As explained earlier in Section 5.3.1 (p. 121), epistemic beliefs include two disparate facets, namely, perceived reliability of knowledge, and perceived justification for knowing (Kammerer et al., 2013). These were used separately to categorize the participants as either epistemologically naïve or epistemologically robust.

The test for moderation was conducted by employing a multi-group PLS analysis (Casaló et al., 2011; Velayutham et al., 2012). As explained earlier in Section 5.4.1 (p. 130), individuals within $\pm 1/2$ standard deviation around the mean of their responses to the questionnaire items measuring epistemic beliefs were eliminated. This was needed to confine the scope of the analysis to participants who were clearly epistemologically naïve or robust (Casaló et al., 2011; Rodríguez et al., 2008). Age group, gender, educational/employment status, Internet anxiety, familiarity with review websites, and attitude toward reviews as well as the website versions to which participants were assigned were included as control variables. None were found to be significant.

6.3.1. Perceived Reliability of Knowledge

One facet of epistemic beliefs is perceived reliability of knowledge. On this facet, the participants were categorized as either epistemologically naïve or robust based on their responses to the questionnaire items measuring the construct—EpiR1 to EpiR11

shown in Table 6.3 (p. 147). Among the participants, 138 were epistemologically naïve while 107 were epistemologically robust.

The participants who were epistemologically naïve with respect to perceived reliability of knowledge were 67.98% accurate in identifying authentic reviews, and 34.60% accurate in identifying fictitious entries. Their overall accuracy was 50.97%. In contrast, the epistemologically robust participants were 70.81% accurate in identifying authentic reviews, and 54.38% accurate in identifying fictitious entries. Their overall accuracy was 62.62%.

Even though the epistemologically robust participants were relatively better than the epistemologically naïve individuals, the performance of the former was not too promising in identifying fictitious reviews (Bond & DePaulo, 2006; Vrij & Baxter, 1999). It was only slightly better than chance. Both epistemologically naïve and robust participants were better in identifying authentic reviews than in spotting fictitious ones. This could be because of humans' inherent truth bias that leads them to perceive most information as being authentic (Vrij & Baxter, 1999). If humans label most entries in a pool of authentic and fictitious reviews as authentic, the individuals' performance in identifying authentic reviews would be conceivably good. This suggests that regardless of epistemic beliefs with respect to perceived reliability of knowledge, the participants' ability to discern authenticity was impaired by truth bias.

Among the participants who were epistemologically naïve with respect to perceived reliability of knowledge ($R^2 = 42.20\%$, $f^2 = 0.73$, $Q^2 > 0$), perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.37$, $t = 3.40$, $p < 0.001$) while perceived specificity showed a positive association ($\beta = 0.38$, $t = 3.53$, $p <$

0.001). However, perceived comprehensibility and perceived tentativeness were not significantly related to perceived authenticity.

Similarly, among the participants who were epistemologically robust with respect to perceived reliability of knowledge ($R^2 = 36.20\%$, $f^2 = 0.57$, $Q^2 > 0$), perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.40$, $t = 3.18$, $p < 0.01$) while perceived specificity showed a positive association ($\beta = 0.32$, $t = 2.74$, $p < 0.01$). The roles played by perceived comprehensibility and perceived tentativeness were non-significant. These results are depicted in Figure 6.3.

Epistemic beliefs with respect to perceived reliability of knowledge did not moderate the relation between any of the four perceived linguistic cues and perceived authenticity. Hence, the proposition P9 could not be supported (cf. Table 2.4, p. 36).

Note. *** represents statistically significant relation at the $p < 0.001$ level.
 ** represents statistically significant relation at the $p < 0.01$ level.
 For epistemologically naïve: sample size = 138, number of data points = 414 (138 x 3).
 For epistemologically robust: sample size = 107, number of data points = 321 (107 x 3).

Figure 6.3: Multi-group PLS result for perceived reliability of knowledge.

6.3.2. *Perceived Justification for Knowing*

The other facet of epistemic beliefs is perceived justification for knowing. On this facet, the participants were categorized as either epistemologically naïve or robust based on their responses to the questionnaire items measuring the construct—EpiJ1 to EpiJ4 shown in Table 6.3 (p. 147). Among the participants, 89 were epistemologically naïve while 117 were epistemologically robust.

On the one hand, the epistemologically naïve participants with respect to perceived justification for knowing were 66.67% accurate in identifying authentic reviews, and 40.15% accurate in identifying fictitious entries. Their overall accuracy was

53.56%. On the other hand, the epistemologically robust participants were 75.15% accurate in identifying authentic reviews, and 45.60% accurate in identifying fictitious entries. Their overall accuracy was 59.83%.

Similar to the results reported in Section 6.3.1 (p. 154), the overall accuracy in discerning review authenticity was higher among the epistemologically robust participants with respect to perceived justification for knowing compared with the epistemologically naïve individuals. Nonetheless, the performance of the epistemologically robust participants was still only marginally better than chance (Bond & DePaulo, 2006; Vrij & Baxter, 1999). Moreover, regardless of epistemic beliefs with respect to justification for knowing, the participants were better in identifying authentic reviews than in spotting fictitious ones. This suggests that their ability to discern review authenticity was impeded by their inherent truth bias (Vrij & Baxter, 1999).

Among the participants who were epistemologically naïve with respect to perceived justification for knowing ($R^2 = 22.70\%$, $f^2 = 0.29$, $Q^2 > 0$), perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.28$, $t = 2.21$, $p < 0.05$). Perceived comprehensibility, perceived specificity, and perceived tentativeness of reviews were not significantly related to perceived authenticity.

Among the participants who were epistemologically robust with respect to perceived justification for knowing ($R^2 = 49.80\%$, $f^2 = 0.99$, $Q^2 > 0$), perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.45$, $t = 4.61$, $p < 0.001$) while perceived specificity showed a positive association ($\beta = 0.37$, $t = 3.63$, $p < 0.001$). However, perceived comprehensibility and perceived tentativeness were not significantly related to perceived authenticity. These results are depicted in Figure 6.4.

The moderating role played by individuals' epistemic beliefs with respect to perceived justification for knowing is summarized in Table 6.7. Specifically, the extent to which perceived justification for knowing moderated the relation between perceived exaggeration and perceived authenticity as well as that between perceived specificity and perceived authenticity were examined. The relation between perceived comprehensibility and perceived authenticity as well as that between perceived tentativeness and perceived authenticity were statistically non-significant for both naïve and robust participants. Hence, examining the moderating effect for these two relations would not have been meaningful.

Note. *** represents statistically significant relation at the $p < 0.001$ level.

* represents statistically significant relation at the $p < 0.05$ level.

For epistemologically naïve: sample size = 89, number of data points = 267 (89 x 3).

For epistemologically robust: sample size = 117, number of data points = 351 (117 x 3).

Figure 6.4: Multi-group PLS result for perceived justification for knowing.

Perceived justification for knowing significantly moderated the relation between perceived exaggeration and perceived authenticity ($t = 18.89$, $p < 0.001$). Specifically, the negative relation between perceived exaggeration and perceived authenticity was stronger for epistemologically robust participants vis-à-vis epistemologically naïve individuals. All else being equal, exaggerated reviews were viewed less favorably by the former in terms of authenticity.

Additionally, perceived justification for knowing significantly moderated the relation between perceived specificity and perceived authenticity ($t = -12.00$, $p < 0.001$). Specifically, the positive relation between perceived specificity and perceived authenticity was stronger for epistemologically robust participants vis-à-vis epistemologically naïve individuals. All else being equal, specific reviews were viewed more favorably by the former in terms of authenticity. These results lend support to the proposition P9 that epistemic beliefs moderate the relations between perceived linguistic cues and perceived authenticity (cf. Table 2.4, p. 36).

Table 6.7: Moderating role played by perceived justification for knowing.

Relation	Epistemologically Naïve		Epistemologically Robust		t-Stat
	β	Std. Error	β	Std. Error	
PExa $\rightarrow$ PA	-0.28	0.13	-0.45	0.10	18.89***
PSpe $\rightarrow$ PA	0.25	0.15	0.37	0.10	-12.00***

Note. PExa: Perceived exaggeration, PSpe: Perceived specificity, PA: Perceived authenticity.

*** represents statistically significant difference between groups at the $p < 0.001$ level.

To summarize, the two facets of epistemic beliefs—perceived reliability of knowledge, and perceived justification for knowing—appear disparate. On the one hand,

epistemic beliefs with respect to perceived reliability of knowledge did not moderate the relation between perceived linguistic cues and perceived authenticity of reviews. On the other hand, epistemic beliefs with respect to perceived justification for knowing significantly moderated the relation between perceived exaggeration and perceived authenticity ($t = 18.89, p < 0.001$) as well as that between perceived specificity and perceived authenticity ($t = -12.00, p < 0.001$). The relations were stronger for the epistemologically robust participants vis-à-vis the epistemologically naïve individuals.

Nonetheless, regardless of epistemic beliefs with respect to either perceived reliability of knowledge or perceived justification for knowing, the participants' performance in discerning review authenticity was not too promising. Their perceptions seemed to be impaired by their inherent truth bias. This could be inferred based on their relatively high accuracy in identifying authentic reviews (ranged from 67.98% to 75.15%), and low accuracy in spotting fictitious entries (ranged from 34.60% to 54.38%). This in turn calls for finding ways to alleviate truth bias among individuals so that they would be cautious in discerning information authenticity.

6.4. Results of the Qualitative Analysis

As explained earlier in Section 5.4.2 (p. 131), the qualitative analysis involved inductive coding of the participants' responses to the open-ended questions (cf. Table 5.4, p. 121). The purpose was to identify the strategies that they used to discern review authenticity.

Five strategies could be identified. These included assessing comprehensibility, checking the level of exaggeration, examining specificity, looking for conflicting

information, and relying on intuition. The coding scheme used to identify these strategies from participants' open-ended responses is shown in Table 6.8. Appendix I (p. 213) presents quotations from the participants related to each of the five strategies.

The first three strategies are consistent with the propositions made in the user framework of perceptions that perceived comprehensibility, perceived exaggeration, and perceived specificity could be related to perceived authenticity. In other words, the first three strategies lend support to the propositions P5, P6 and P7 respectively (cf. Table 2.4, p. 36). The other two strategies were newly uncovered by leveraging on the exploratory nature of the qualitative analysis.

Table 6.8: Coding scheme for the qualitative analysis.

Strategies	Coding Procedure	Examples
Comprehensibility	Responses were coded as comprehensibility-related when they focused on the ease with which reviews could be read or understood.	"too many grammatical errors" (participant 236)
Exaggeration	Responses were coded as exaggeration-related when they focused on the extent to which reviews used superlatives, or went overboard to appear convincing.	"too much exaggerated" (participant 136)
Specificity	Responses were coded as specificity-related when they focused on the informativeness or the richness of specific details in reviews.	"good details provided" (participant 258)
Conflicting information	Responses were coded as conflicting information-related when they focused on inconsistencies in the available information.	"title matched the description" (participant 258)
Intuition	Responses were coded as intuition-related when they focused on gut feeling or instincts.	"seems normal" (participant 28)

The first strategy to discern authenticity involved assessing the comprehensibility of reviews. For example, participant 236 was concerned due to "too many grammatical errors" in a review. Participant 238 found a review to be "very unclear about what the

reviewer wants to say.” Likewise, participant 260 commented, “sentence structure is bad.” On perceiving these comprehensibility-related cues, all these participants ended up labeling authentic reviews as fictitious. Perceived comprehensibility did not seem to help individuals in accurately discerning review authenticity.

Epistemologically naïve participants were poor in accurately discerning review authenticity based on comprehensibility. For example, participant 78, who was epistemologically naïve with respect to reliability of knowledge as well as justification for knowing, inaccurately labeled an authentic review as fictitious because of its “disorganized write up.” Participant 322, another individual with similar epistemic beliefs, inaccurately labeled a fictitious review as authentic because “it stated [its arguments] clearly.”

The performance of epistemologically robust participants in accurately discerning review authenticity based on comprehensibility also resembled random guessing. For example, participant 172, who was epistemologically robust with respect to perceived justification for knowing stated, “The statements are . . . not well written.” Due to the poorly written statements in an authentic review, the participant labeled it as fictitious. Participant 95, who was epistemologically robust with respect to reliability of knowledge as well as justification for knowing, accurately identified a fictitious review because it was perceived as having been written by someone who was not “conversant in English.”

The second strategy to discern authenticity involved checking the level of exaggeration in reviews. For example, participant 151 accurately labeled a review as fictitious because its level of exaggeration made it appear “more like an advertisement.” Participant 136 accurately identified a fictitious review that was perceived as being “too

much exaggerated.” However, individuals’ perceived exaggeration occasionally led them to make inaccurate decisions. For example, participant 310 labeled an authentic review as fictitious because it appeared to make overly “sweeping statements.” This suggests that while fictitious reviews are generally exaggerated, authentic reviews are not always innocuous.

Nonetheless, barring a few exceptions, individuals’ perceptions of the level of exaggeration in reviews generally helped them to accurately discern review authenticity regardless of epistemic beliefs. For example, participant 203, who was epistemologically naïve with respect to reliability of knowledge as well as justification for knowing, accurately identified an authentic review because it was perceived as one providing “a balanced opinion.” Participant 163, another individual with similar epistemic beliefs, accurately identified a fictitious review because it was perceived as being “a bit too exaggerated in tone.” Furthermore, participant 335, who was epistemologically robust with respect to reliability of knowledge as well as justification for knowing, accurately identified an authentic review that provided “both [positive and negative] sides of the story. Did not try to understate or overstate. very balanced.” Participant 266, who was epistemologically robust with respect to justification for knowing, accurately identified a fictitious review that seemed “overly enthusiastic.”

The third strategy to discern authenticity involved examining the specificity of reviews. For example, participant 6 accurately identified an authentic review, “More specific details given in the feedback gives the perception of authenticity.” Likewise, participant 258 accurately identified an authentic review because of the “good details provided by the reviewer.” However, individuals’ perceived specificity occasionally led

them to make inaccurate decisions. For example, participant 28 labeled an authentic review as fictitious because it did not state “many details.” This suggests that not all authentic reviews are necessarily rich in specific details.

Nonetheless, barring a few exceptions, individuals’ perceptions of the level of specificity in reviews generally helped them to accurately discern review authenticity regardless of epistemic beliefs. For example, participant 196, who was epistemologically naïve with respect to reliability of knowledge as well as justification for knowing, accurately labeled an authentic review because it was perceived as being “quite contextual.” Participant 189, another individual with similar epistemic beliefs, accurately identified a fictitious review stating, “The review is very generic and hardly provide any information that gives a sneak peek into the hotel’s environment. I can’t picture the hotel in my mind.” Furthermore, participant 41, who was epistemologically robust with respect to reliability of knowledge as well as justification for knowing, accurately identified a review as authentic because it apparently contained “quite a lot of details with adequate information.” Participant 18, another individual with similar epistemic beliefs, accurately identified a fictitious review because it was perceived to contain “too little details.”

The fourth strategy to discern authenticity involved looking for conflicting information. This strategy did not always help participants in accurately discerning review authenticity. For example, participant 79 labeled an authentic review as fictitious because its sentences “on hotel services” were perceived as being self-contradictory. Participant 312 labeled a fictitious review as authentic after comparing its arguments with those present in the other two available reviews on the LoveToTravel.com version to which the individual was assigned, “What was written differs from the rest of the reviews.”

The performance of epistemologically naïve participants in accurately discerning review authenticity based on conflicting information resembled random guessing. For example, participant 347, who was epistemologically naïve with respect to reliability of knowledge as well as justification for knowing, inaccurately labeled an authentic review as fictitious. The argument was that the review did not “match the ratings.” Participant 322, another individual with similar epistemic beliefs, accurately labeled a fictitious review by comparing it with the other two entries available on the LoveToTravel.com version to which the individual was assigned, “It does not even mention about the cleanliness of the room like the other 2 reviews. If it is authentic, the poster should . . . mention about it.”

Epistemologically robust participants often succeeded in accurately discerning review authenticity by looking for conflicting information. For example, participant 258, who was epistemologically robust only with respect to justification for knowing, accurately identified an authentic review because its title “matched with the description.” Participant 43 who was epistemologically robust with respect to reliability of knowledge as well as justification for knowing commented about a fictitious review, “The person complains about wifi and then replies ‘who needs to pay for wifi in hotels nowadays’. There is a contradiction.” The perceived contradiction prompted the participant to accurately label it as fictitious.

The fifth strategy to discern authenticity involved relying on intuition. For example, participant 28 identified an authentic review accurately by arguing, “[it] seems normal.” Likewise, participant 168 identified a fictitious review accurately indicating,

“[it] seems fishy.” Conceivably, such gut feeling might not help individuals to accurately discern review authenticity.

Regardless of epistemic beliefs, participants’ ability to discern review authenticity by relying on intuition resembled random guessing. For example, participant 196, who was epistemologically naïve with respect to reliability of knowledge as justification for knowing, accurately labeled an authentic review because it “felt authentic.” Participant 331, another individual with similar epistemic beliefs, failed to identify a fictitious review, “seems not to be a fake one.” Again, participant 54, who was epistemologically robust with respect to reliability of knowledge as justification for knowing, accurately identified an authentic review by indicating, “sounds quite believable.” However, participant 227, an individual with similar epistemic beliefs, inaccurately labeled a fictitious review as authentic by stating, “seems like a believable story.”

Interestingly, there was no clear majority in the use of the five strategies. However, most participants used only one of the five strategies to decide whether a given review was authentic or fictitious. Inability to employ multiple strategies per review points to the limited cognitive efforts that humans make to discern authenticity.

Moreover, majority of the participants employed one of these five strategies per review only on descriptions, and ignored titles of the entries. For example, participant 166 indicated, “The review titles do not determine the authenticity . . . their purpose is to summarize the descriptions of the reviews.” Participant 180 found that it was “too hard to guess [review authenticity] from the title.”

A few participants, most of whom were epistemologically naïve with respect to reliability of knowledge as well as justification for knowing (e.g., participant 11,

participant 328), examined neither titles nor descriptions to discern review authenticity. They were completely dependent on intuition.

Table 6.9 summarizes the quantitative as well as the qualitative results pertaining to the framework of user perceptions. Even though the quantitative analysis did not find evidence of perceived comprehensibility to be related to perceived authenticity, the qualitative analysis found that the participants occasionally relied on comprehensibility to discern whether reviews were authentic or fictitious. Hence, the proposition P5 could be partially supported (cf. Table 2.4, p. 36).

According to the quantitative analysis, perceived exaggeration was negatively related to perceived authenticity ($\beta = -0.39$, $t = 3.69$, $p < 0.001$) while perceived specificity showed a positive association ($\beta = 0.34$, $t = 3.00$, $p < 0.01$). The qualitative analysis also found that the participants checked the level of exaggeration, and examined specificity in reviews to discern their authenticity. These results lend support to the propositions P6 and P7 (cf. Table 2.4, p. 36).

The quantitative analysis did not find perceived tentativeness to be significantly related to perceived authenticity of reviews. The qualitative analysis further confirmed that the participants did not rely on the tentativeness of reviews to determine if the entries were authentic or fictitious. Hence, the proposition P8 could not be supported (cf. Table 2.4, p. 36).

The quantitative analysis found epistemic beliefs with respect to perceived justification for knowing to significantly moderate the relation between perceived exaggeration and perceived authenticity ($t = 18.89$, $p < 0.001$) as well as that between perceived specificity and perceived authenticity ($t = -12.00$, $p < 0.001$). The qualitative

analysis suggested that epistemologically robust participants could be more cautious in discerning review authenticity compared with epistemologically naïve individuals. However, according to the quantitative analysis, epistemic beliefs with respect to perceived reliability of knowledge did not moderate the relation between perceived linguistic cues and perceived authenticity. Hence, H9 could only be partially supported (cf. Table 2.4, p. 36).

Additionally, the qualitative analysis uncovered two new strategies that participants employed to discern review authenticity. These included looking for conflicting information, and relying on intuition.

Table 6.9: Summary of the results pertaining to the framework of user perceptions.

Propositions	Findings	Additional Insights
P5: Perceived comprehensibility is related to perceived review authenticity.	Partially supported	Perceived exaggeration was negatively related to perceived review authenticity.
P6: Perceived exaggeration is related to perceived review authenticity.	Supported	Perceived specificity was positively related to perceived review authenticity.
P7: Perceived specificity is related to perceived review authenticity.	Supported	Epistemic beliefs only with respect to justification for knowing moderated the relation between perceived exaggeration and perceived authenticity as well as that between perceived specificity and perceived authenticity.
P8: Perceived tentativeness is related to perceived review authenticity.	Not supported	Individuals often look for conflicting information, and rely on intuition to discern review authenticity.
P9: Epistemic beliefs moderate the relation between perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity as well as perceived tentativeness—and perceived review authenticity.	Partially supported	The relation between perceived comprehensibility and perceived authenticity was supported by the qualitative analysis but not by the quantitative analysis.

6.5. Results of the Robustness Check

Overall, human ability to discern authenticity of reviews does not appear too promising. For one, their decision-making seems to be adversely affected by their inherent truth bias. Their performance in discerning review authenticity resembles shots in the dark. Moreover, they rarely use multiple strategies to determine if a given review is authentic or fictitious. Furthermore, they seem to rely solely on descriptions but ignore titles in determining review authenticity. These justify the need to guide humans how to discern review authenticity—a purpose that is served in this research by the Robustness Check.

The Robustness Check was specifically needed to wrap up this research because on the one hand, the Linguistic Study suggested that automated techniques that rely on linguistic analysis help ascertain review authenticity (cf. Table 4.11, p. 96). On the other hand, the User Study revealed that humans often fail to accurately discern review authenticity regardless of their epistemic beliefs (cf. Table 6.9, p. 169).

Concurrently, the literature suggests that guidance on ways to evaluate information improves individuals' information-processing strategies (Argelagós & Pifarré, 2012; Munzel, 2015). This sets an interesting context to examine if the actual differences between authentic and fictitious reviews found in the Linguistic Study could be used to teach humans how to discern review authenticity.

For this purpose, a guideline was developed as explained earlier in Section 5.5 (p. 134). To test its efficacy, the guideline was used as an intervention in a between-participants experimental design that involved 240 annotators. Half of the annotators formed the control group while the other half constituted the treatment group. Those in the

control group were required to predict authenticity of a set of 54 reviews (cf. Section 5.1.2, p. 110)—27 authentic and the other 27 fictitious arranged in a random sequence—on their own. Those in the treatment group were required to predict authenticity of the reviews using the guideline that served as an intervention.

For all annotators, the proportion of reviews whose authenticity was accurately identified was checked. The accuracy percentage of the control group (54.32 ± 7.98) was significantly inferior compared with that of the treatment group (68.94 ± 7.23); $t(238) = -14.86$, $p < 0.001$.

To delve deeper, the volumes of accurately identified authentic reviews as well as accurately identified fictitious reviews were checked. In accurately identifying the 27 authentic reviews, the control group (19.19 ± 4.67) appeared slightly better than the treatment group (18.10 ± 4.75). No significant difference was however found in terms of accuracy percentage between the former (71.08 ± 17.28) and the latter (67.04 ± 17.60); $t(238) = 1.80$, $p = 0.07$.

In accurately identifying the 27 fictitious reviews, the performance of the control group (10.14 ± 3.92) was worse than that of the treatment group (19.13 ± 3.88). The accuracy percentage of the former (37.56 ± 14.50) was significantly inferior compared with that of the later (70.83 ± 14.38); $t(238) = -17.85$, $p < 0.001$. This finding demonstrates that the guideline developed in this research improved annotators' ability to identify fictitious reviews.

Interestingly however, the control group had higher accuracy in identifying authentic reviews vis-à-vis the treatment group albeit statistically non-significant. A possible explanation lies in truth bias (Vrij & Baxter, 1999), which perhaps led annotators

in the control group to annotate most reviews as authentic. If they blindly label most of the reviews provided to them as authentic, they can certainly perform well in identifying authentic reviews.

To verify if truth bias is a valid explanation, the volumes of reviews annotated as authentic by the two groups of annotators was further checked. Indeed, the control group (36.05 ± 7.46) annotated significantly more reviews as authentic compared with the treatment group (25.98 ± 7.75); $t(238) = 10.26$, $p < 0.001$. This indicates that the guideline not only improved annotators' ability to identify fictitious reviews but also reduced their inherent truth bias, which is one of the biggest impediments for humans in discerning the authenticity of information (Bond & DePaulo, 2006; Vrij & Baxter, 1999).

6.6. Comparing User Study Results with Prior Works

The results of the User Study show that individuals' perceived linguistic cues could explain their perceptions of review authenticity reasonably well. Epistemic beliefs with respect to perceived justification for knowing moderated the relation between perceived linguistic cues—particularly, perceived exaggeration as well as perceived specificity—and perceived authenticity. These results generally support the propositions of the framework of user perceptions.

Six main findings were gleaned from the results. First, the participants' perceptions of exaggeration and specificity in reviews seemed to help discern authenticity of the entries. As hinted in prior works (Connors et al., 2011; Papathanassis & Knolle, 2011), the quantitative analysis found that reviews that are specific but with little exaggeration stand a good chance to be deemed as being authentic by the participants. In

addition, the qualitative analysis further confirmed that individuals' perceptions of exaggeration as well as specificity in reviews often helped them accurately distinguish between authentic and fictitious reviews. This demonstrates the crucial role played by the two linguistic cues—exaggeration and specificity—in helping individuals discern review authenticity.

Second, the participants' perceptions of comprehensibility and tentativeness in reviews did not help discern authenticity of the entries. In contrast to prior works (Ekman & Friesen, 1969; Latta & Buck, 2008; Porter et al., 2012; Schrank et al., 2010), the participants did not deem reviews that were comprehensible but with little tentativeness as being authentic. That is why the quantitative analysis could not detect any significant relation between perceived comprehensibility and perceived authenticity as well as that between perceived tentativeness and perceived authenticity. The qualitative analysis revealed that the participants occasionally assessed the comprehensibility of reviews. However, even after perceiving such comprehensibility-related cues, several participants ended up discerning review authenticity incorrectly. This finding contradicts the notion that alertness toward the content of information helps distinguish between truth and fiction (Sperber et al., 2010).

Third, most of the participants did not rely on titles of reviews to discern authenticity. They were mostly inclined to examine only descriptions even though reviews are commonly structured as titles and descriptions. This explains why questionnaire items that were specifically related to titles (e.g., PCom1, PSpe1, PExa1, and PTen1) had non-significant item weightings as shown earlier in Table 6.4 (p. 149). Titles of reviews are known to play crucial communicative roles (Ascaniis & Gretzel, 2012). They are often

displayed more conspicuously on review websites than descriptions (Ludwig et al., 2013). Yet, individuals appear unapprised about the purpose served by titles in reviews.

Fourth, the qualitative analysis revealed two serendipitous strategies, namely, looking for conflicting information, and relying on intuition, that were used by the participants to discern authenticity. These two strategies have not received much attention in the literature. Yet, they were empirically uncovered through the qualitative analysis. The use of these strategies notwithstanding, the participants' performance in discerning review authenticity resembled shots in the dark. This offers a possible explanation on why previous works found human ability to distinguish between truth and fiction to resemble random guessing (Bond & DePaulo, 2006; Porter & ten Brinke, 2008). Nonetheless, this finding offers the possibility to extend previous works on linguistic differences between authentic and fictitious reviews (Ott et al., 2011; Harris, 2012) by taking into account conflicting information between titles and descriptions of a given review as well as conflicting information among multiple reviews.

Fifth, epistemic beliefs with respect to perceived justification for knowing moderated the relation between perceived linguistic cues and perceived authenticity. Specifically, exaggerated reviews were viewed less favorably by epistemologically robust participants vis-à-vis epistemologically naïve individuals in terms of authenticity. Moreover, specific reviews were viewed more favorably by the former in terms of authenticity. However, epistemic beliefs with respect to perceived reliability of knowledge did not emerge as a significant moderator. This finding confirms previous works that suggested that epistemic beliefs do not constitute an atomic construct (Hofer, 2004; Hofer & Pintrich, 1997). Rather, epistemic beliefs encompass two disparate

constructs, namely, perceived reliability of knowledge, and perceived justification for knowing (Kammerer et al., 2013). Furthermore, this finding extends the previous works (Mason & Ariasi, 2010; Whitmire, 2004) by highlighting that perceived justification for knowing could be more influential than perceived reliability of knowledge in affecting the relation between individuals' perceived linguistic cues and perceived authenticity. Nonetheless, in terms of ability to discern authenticity, the accuracy was slightly higher among the epistemologically robust participants (with respect to perceived reliability of knowledge, overall accuracy = 62.62%, with respect to perceived justification for knowing, overall accuracy = 59.83%) vis-à-vis the epistemologically naïve individuals (with respect to perceived reliability of knowledge, overall accuracy = 50.97%, with respect to perceived justification for knowing, overall accuracy = 53.56%) as indicated earlier in Section 6.3 (p. 154).

Finally, the guideline developed for the purpose of the Robustness Check improved human ability to discern review authenticity by lowering individuals' inherent truth bias. Compared with previous works such as Wiley et al. (2009) that used a one-hour long instructional unit as a guideline, the one used here was much shorter (cf. Appendix G, p. 207). Even then, it was efficacious in improving human ability to identify fictitious reviews amid authentic ones. This encouraging finding lends support to the premise that interventions on critical evaluation of information improves humans' information-processing strategies (Argelagós & Pifarré, 2012). It is high time to use similar training materials to develop individuals' online information literacy skills (Gross & Latham, 2012).

6.7. Summary of the User Study – Findings

This chapter has presented the findings of the User Study. Specifically, it began by presenting the descriptive statistics of the dataset. Next, it presented the results of RO 2.1, which examined the relation between perceived linguistic cues and perceived authenticity of reviews. Perceived exaggeration was negatively related to perceived authenticity. Moreover, perceived specificity was positively related to perceived authenticity.

Thereafter, this chapter presented the results of RO 2.2, which focused on the moderating role of epistemic beliefs. Epistemic beliefs with respect to perceived justification for knowing significantly moderated the relation between perceived exaggeration and perceived authenticity as well as that between perceived specificity and perceived authenticity. Specifically, exaggerated reviews were viewed less favorably by epistemologically robust participants vis-à-vis epistemologically naïve individuals in terms of authenticity. Moreover, specific reviews were viewed more favorably by the former in terms of authenticity. The results however did not paint an overly promising picture about human ability to discern review authenticity regardless of epistemic beliefs.

Next, this chapter presented the results of the Robustness Check, which investigated the extent to which humans could be taught how to discern review authenticity. The guideline that was developed for this purpose improved human ability to identify fictitious reviews. It also reduced individuals' default tendency to label every review as authentic. The results have been discussed in light of the literature.

Having presented the findings of the User Study, this chapter has laid the foundation to bring the thesis to a close. Hence, the next chapter is devoted to concluding the thesis.

CHAPTER SEVEN CONCLUSION

This chapter concludes the research on authentic and fictitious online reviews. As indicated earlier in Section 1.1.1 (p. 1), even though fictitious reviews are widespread on the Internet, they are not easily distinguishable from authentic ones. To address the problem, this research investigated the ways in which linguistic differences between authentic and fictitious reviews were related to users' perceived review authenticity. Two research questions (RQs) were formulated. The first RQ investigated actual linguistic differences between authentic and fictitious reviews while the second RQ investigated users' perceived authenticity of reviews (cf. Figure 1.1, p. 9).

To address the RQs, this research theorized how the linguistic differences between authentic and fictitious reviews could be related to users' perception of review authenticity. In so doing, the *theoretical model of authentic and fictitious reviews (TAF)* was developed in Section 2.3 (p. 23) based on the theoretical perspectives of deception and information processing. Moving down the ladder of abstraction, the TAF was conceived as the *framework of linguistic cues* and the *framework of user perceptions*. The former undergirded RQ 1 while the latter underpinned RQ 2. Two studies were proposed such that the first (Linguistic Study) addressed RQ 1, and the second (User Study) addressed RQ 2. To further confirm the findings arising from the two studies, an additional investigation (Robustness Check) was conducted.

The Linguistic Study investigated linguistic differences between authentic and fictitious reviews. As described earlier in Section 3.3 (p. 45), a dataset of authentic and fictitious reviews was collected. Differences between authentic and fictitious reviews

were analyzed using classification techniques, followed by feature selection, independent samples t-tests as well as word-level analysis.

The User Study investigated users' perceived authenticity of reviews. As described earlier in Section 5.2 (p. 112), data were collected from participants using an online survey. Data were analyzed using partial least squares structural equation modeling, and content analysis. Finally, the Robustness Check investigated the possibility of developing a guideline to help humans identify the linguistic differences between authentic and fictitious reviews using an experimental design.

As shown in Figure 7.1, the remainder of the chapter is organized into five sections. The first section (p. 179) summarizes the research findings. The second section (p. 182) highlights the contributions of this research. Both theoretical contributions (p. 182), and practical contributions (p. 189) are teased out. The third section (p. 193) indicates the ethical considerations involved in this research. The fourth section (p. 194) acknowledges the limitations of this research. Finally, the fifth section (p. 195) offers directions for future research.

Figure 7.1: Overview of Chapter Seven.

7.1. Summary of the Research Findings

The goal of this research was to investigate the ways in which linguistic differences between authentic and fictitious reviews are related to users' perceived review authenticity. To accomplish the research goal, two RQs were formulated, each of which was expressed as two subordinate research objectives (ROs). The findings are summarized below.

RQ 1: *How do authentic reviews linguistically differ from fictitious entries?*

RO 1.1: To distinguish between authentic and fictitious reviews based on linguistic cues.

RO 1.2: To identify specific linguistic traits unique to authentic and fictitious reviews.

With respect to RO 1.1, this research identified four linguistic cues to distinguish between authentic and fictitious reviews. These included comprehensibility, exaggeration,

specificity and tentativeness. Together, they were operationalized as 83 variables that were used as features for classification. These features outperformed the two selected baselines—the LIWC-baseline and the Bigrams-baseline—in distinguishing between authentic and fictitious reviews (cf. Section 4.2, p. 80).

With respect to RO 1.2, this research identified unique linguistic traits of authentic and fictitious reviews corresponding to 38 of the 83 variables. These variables straddled across the four linguistic cues of comprehensibility (e.g., long words), exaggeration (e.g., emotiveness), specificity (e.g., nouns), and tentativeness (e.g., fillers). Specifically, there were 19 exaggeration-related, nine specificity-related, five comprehensibility-related, and five tentativeness-related linguistic traits (cf. Section 4.3, p. 81).

The findings of RO 1.1 and RO 1.2 can be lifted to answer RQ 1 as follows: Authentic reviews are linguistically different from fictitious entries. Specifically, exaggeration offers maximal opportunity to distinguish between authentic and fictitious reviews. It is followed by specificity. On the other hand, comprehensibility and tentativeness offer relatively less opportunity to ascertain review authenticity.

RQ 2: To what extent are users able to perceive the authenticity of reviews?

RO 2.1: To analyze how perceived linguistic cues are related to perceived authenticity of reviews.

RO 2.2: To examine how epistemic beliefs affect the relation between perceived linguistic cues and perceived authenticity of reviews.

With respect to RO 2.1, users' perceived linguistic cues could explain perceived review authenticity reasonably well. Specifically, perceived exaggeration was negatively

related to perceived authenticity. In contrast, perceived specificity was positively related to perceived authenticity (cf. Section 6.2, p. 151).

With respect to RO 2.2, epistemic beliefs with respect to perceived justification for knowing significantly moderated the relation between perceived exaggeration and perceived authenticity as well as that between perceived specificity and perceived authenticity. Exaggerated reviews were viewed less favorably in terms of authenticity by epistemologically robust individuals vis-à-vis epistemologically naïve ones. Conversely, specific reviews were viewed more favorably by the former. Overall, regardless of epistemic beliefs, individuals' ability to predict if a review was authentic or fictitious was not too good. Their accuracy was found to be only marginally better than chance (cf. Section 6.3, p. 154).

The findings of RO 2.1 and RO 2.2 can be lifted to answer RQ 2 as follows: Users' perceptions of exaggeration and specificity in reviews are related to the ways in which they perceive the authenticity of the entries. The relations are stronger among users who are epistemologically robust with respect to perceived justification for knowing vis-à-vis epistemologically naïve individuals. However, human ability to discern review authenticity does not appear too promising regardless of epistemic beliefs.

By answering the RQs, this research finds that on the one hand, language of reviews helps ascertain their authenticity. On the other hand, humans are not sufficiently adept in discerning review authenticity. Nonetheless, the language-based guideline that was developed in this research (cf. Appendix G, p. 207) significantly improved human ability to identify fictitious reviews amid authentic ones.

7.2. Contributions

7.2.1. Theoretical Contributions

This research makes six theoretical contributions. First, it advances the literature by developing the theoretical model of authentic and fictitious reviews (TAF). This is a significant contribution because developing theories is considered as the jewel in the crown of scholarly efforts (Eisenhardt, 1989). Theories are important as they enable systematic accumulation of knowledge that in turn enlightens professional practice (Gregor, 2006). Despite several works done in recent years on the area of authentic and fictitious reviews (Harris, 2012; Ott et al., 2013), this research represents one of the earliest steps toward distilling the essence of existing works in a coherent and parsimonious format to theorize the linguistic differences—both actual and perceived—between authentic and fictitious reviews.

Specifically, the TAF posited that four linguistic cues—comprehensibility, exaggeration, specificity and tentativeness—help predict the actual authenticity of reviews. This part of the TAF that drew from the deception literature was referred as the framework of linguistic cues. Furthermore, the TAF posited that users' perceptions of the same four linguistic cues—in other words, perceived comprehensibility, perceived exaggeration, perceived specificity, and perceived tentativeness—are related to perceived authenticity of reviews albeit differently for epistemologically naïve and robust individuals. This part of the TAF that drew from the information processing literature was referred as the framework of user perceptions. Table 7.1 along with Figure 7.2 summarize the results pertaining to the TAF after empirical validation.

Table 7.1: Summary of the results pertaining to the TAF.

	Propositions of the TAF	Findings	Additional Insights
Framework of Linguistic Cues	P1: Comprehensibility helps ascertain review authenticity.	Supported	Comprehensibility, exaggeration, specificity and tentativeness cumulatively outperformed the LIWC-baseline and the Bigrams-baseline in ascertaining review authenticity.
	P2: Exaggeration helps ascertain review authenticity.	Supported	
	P3: Specificity helps ascertain review authenticity.	Supported	Exaggeration offered maximal opportunity to ascertain review authenticity, followed by specificity.
	P4: Tentativeness helps ascertain review authenticity.	Supported	Compared with exaggeration and tentativeness, comprehensibility and tentativeness offered relatively less opportunity to ascertain review authenticity.
Framework of User Perceptions	P5: Perceived comprehensibility is related to perceived review authenticity.	Partially supported	Perceived exaggeration was negatively related to perceived review authenticity.
	P6: Perceived exaggeration is related to perceived review authenticity.	Supported	Perceived specificity was positively related to perceived review authenticity.
	P7: Perceived specificity is related to perceived review authenticity.	Supported	Epistemic beliefs only with respect to justification for knowing moderated the relation between perceived exaggeration and perceived authenticity as well as that between perceived specificity and perceived authenticity.
	P8: Perceived tentativeness is related to perceived review authenticity.	Not supported	Individuals often look for conflicting information, and rely on intuition to discern review authenticity.
	P9: Epistemic beliefs moderate the relation between perceived linguistic cues—perceived comprehensibility, perceived exaggeration, perceived specificity as well as perceived tentativeness—and perceived review authenticity.	Partially supported	The relation between perceived comprehensibility and perceived authenticity was supported by the qualitative analysis but not by the quantitative analysis.

Figure 7.2: Diagrammatic representation of the results of the TAF.

The second theoretical contribution of this research is the amalgamation of compartmentalized bodies of knowledge. As highlighted earlier in Section 1.1.2 (p. 3), previous works on review authenticity can be grouped into two research clusters. Works in one cluster analyze differences between authentic and fictitious reviews but ignore

users' perceptions. These works have a technical orientation, and are generally conducted by computer science scholars using classification algorithms (Abulaish & Bhat, 2015; Jindal & Liu, 2008). Conversely, works in the other cluster analyze users' perceptions of reviews yet fail to uncover whether users are able to discern review authenticity in the first place. These works have a socio-psychological orientation, and are generally conducted by social scientists using user studies (Gupta & Harris, 2010; Qiu et al., 2012; Sidali et al., 2009). Few works hitherto have attempted to bridge the chasm between these two research clusters.

To address the piecemeal scholarship, this research envisioned the two existing bodies of works as mosaic pieces that could synergize into a more effective way to tackle the theme of review authenticity. It initially distinguished between authentic and fictitious reviews from a technical perspective using classification. Next, this research examined human ability to discern review authenticity from a socio-psychological perspective using a survey followed by an experiment. Specifically, in the Robustness Check, this research fed the results of a technical study as inputs for a socio-psychological study. Such a methodological symbiosis has been seldom exploited. Going forward, this research serves as a call for scholars to explore the possibilities of integrating research methods from different paradigms in order to address research issues at stake.

The third theoretical contribution of this research lies in conceptualizing authentic and fictitious reviews by widening their scope in terms of hotel categories, review polarities as well as review structure. Prior works have mostly confined the scope of their investigation to reviews for popular hotels, and specifically, luxury properties. For example, Yoo and Gretzel (2009) analyzed reviews for only a luxury hotel while Ott et al.

(2011) studied entries for highly rated hotels in TripAdvisor.com. Moreover, some works limited their datasets to only positive reviews (Yoo & Gretzel, 2009), or negative reviews (Ott et al., 2013). Little scholarly attention has been given to the authenticity of moderate reviews. Additionally, the investigation of reviews has been confined to only descriptions of the entries (Harris, 2012; Ott et al., 2011) even though popular review websites such as Amazon.com, IMDb.com and TripAdvisor.com seek entries that are structured in the form of titles followed by descriptions. In fact, titles are generally displayed more conspicuously, and hence, might command greater attention than descriptions of reviews (Ascaniis & Gretzel, 2012; Cao et al., 2011; Ludwig et al., 2013).

Therefore, this research dovetails previous works by widening the scope of investigation to reviews submitted in evaluation of three categories of hotels—luxury, budget and mid-range. Moreover, the reviews uniformly straddled across three polarities: positive, negative and moderate. Additionally, it analyzed both titles and descriptions of reviews. In this way, this research reminds scholars to emphasize on a wider conceptualization of research issues to enhance the breadth of applicability of findings.

The fourth theoretical contribution of this research is a methodological template to develop ground truth, which is indispensable for research on deception (Mihalcea & Burzo, 2012). An elusive problem for scholars working on deception stems from the difficulty in creating ground truth. Specifically, the literature lacks a sound approach to ensure the validity of ground truth data.

In this vein, this research demonstrates that valid ground truth could be developed through three steps. The first step involves identifying appropriate data sources that facilitate collecting authentic information. The sources must be unlikely to be infiltrated

(Ong et al., 2014). For this reason, as justified earlier in Section 3.2.1 (p. 42), this research collected authentic reviews from authenticated review websites such as Agoda.com, Expedia.com and Hotels.com, which allow reviewers to post entries only if they had made a booking to stay in the hotel.

The second step involves collecting authentic information from contributors, who had disclosed maximal information about themselves. Having details about the profile of those who had contributed authentic information could be used to maximize comparability with the background of those who submit fictitious information. For this reason, as explained earlier in Section 3.3.1 (p. 46), this research collected authentic reviews from the authenticated review websites only if reviews were accompanied with the metadata of reviewers' country of origin. Given that the selected review websites do not always allow reviewers to disclose their age and educational profile, the literature was utilized to approximate such reviewer-related demographics information (cf. Section 3.1, p. 38).

The third step involves soliciting fictitious information from participants in a research setting. The background of the recruited participants should be similar to that of individuals who contribute authentic information. Moreover, they must be capable of performing the necessary fictitious behavior. Adequate efforts need to be invested to ensure comparable quantity of fictitious information against the volume of authentic information across profile of the contributors. Furthermore, participants should not be pressurized to create fictitious information. If participants are pushed, they might create fictitious information perfunctorily and spontaneously, thereby leaking more linguistic cues for detection (Bond & DePaulo, 2006; DePaulo et al., 2003; Vrij, 2008). This in turn would take a toll on the quality of the ground truth dataset. For these reasons, as indicated

earlier in Section 3.1 (p. 38), participants were recruited to write fictitious reviews in this research only if they were acquainted with the use of reviews. Furthermore, care was taken to ensure comparable number of authentic and fictitious reviews across profile of contributors as much as possible. Additionally, as indicated earlier in Section 3.3.2 (p. 50), participants were not provided stringent deadlines to write fictitious reviews.

The fifth theoretical contribution of this research is a holistic operationalization of the four linguistic cues, namely, comprehensibility, exaggeration, specificity and tentativeness (cf. Table 3.5, p. 64). This is a significant contribution because scholars strongly recommend operationalizing constructs holistically coupled with identification of their underlying dimensions for precise empirical examination (McLeod & Pan, 2005).

The operationalization of some of these linguistic cues has thus far been largely inconsistent. For example, works such as Yoo and Gretzel (2009) operationalized comprehensibility as average length of words but overlooked indicators of readability. Works such as Harris (2012) operationalized the same construct using a readability indicator without taking into account aspects such as word familiarity. Lack of scholarly consensus about the operationalization of linguistic cues could be a factor that has hitherto hindered development of theories in research on online review authenticity (McLeod & Pan, 2005; Hauch et al., 2015). This research dovetails the literature by operationalizing the linguistic cues in terms of variables such as lexical diversity, and the use of ellipses, which have at best received cursory attention in previous studies. This operationalization could inform future research on text analysis and mining.

The final theoretical contribution of this research stems from extending the scholarly understanding of the role of guidelines in information processing. The finding

that the guideline (Appendix G, p. 207) improved human ability to discern review authenticity demonstrates the importance of training to address the information-seeking problems of “discerning information from misinformation [or disinformation], screening information for its quality, and assessing the credibility of its resources” (Rubin & Lukoianova, 2015, p. 916). Specifically, this research found that a guideline could not only improve human ability to discern review authenticity but also enhance immunity against their inherent truth bias, which is regarded as one of the biggest impediments for individuals in distinguishing between truth and fiction (Vrij & Baxter, 1999; Vrij, 2008).

Currently, users are often warned by review websites about the possible prevalence of fictitious entries. Yet, they receive little guidance on ways to identify fictitious reviews amid authentic ones. Given that authenticity of user-generated content—and not specifically reviews—on social media applications is a growing concern, it is high time that scholars expend more efforts on systematic development of guidelines to help sharpen the information-processing strategies of Internet users (Watson, 2014). By shifting the scholarly focus toward development of training materials for information processing, this research hopes to pave the way for better online experiences for Internet users in the future.

7.2.2. Practical Contributions

This research offers several practical implications. For users who read reviews on the Internet, it offers a note of caution. On the one hand, not all reviews are guaranteed to be authentic. On the other, fictitious reviews could be written to closely mimic authentic ones. As a result, users run the risk of being deceived by fictitious reviews in making

purchase decisions. Nonetheless, they could lean on the guideline developed in this research (Appendix G, p. 207) to discern review authenticity with reasonable accuracy. Additionally, this research recommends users to cultivate robust epistemic beliefs that would make them more cautious in information processing tasks.

For users who contribute reviews on the Internet, this research serves to remind the value of netiquette by highlighting the menace of fictitious entries (Buelens et al., 2007). After all, individuals who contribute reviews might themselves need to look for entries in order to make purchase decisions. Under such circumstances, they would stand a good chance to get a taste of their own medicine. If users do not share honest post-purchase experiences, the *raison d'être* of review websites would be called into question.

Besides, this research has the potential to not only help identify businesses that maliciously hype their own offerings but also alert businesses whose offerings had been slandered by their potential competitors. Hence, if businesses find suspicious-looking reviews posted consistently on their offerings, this research recommends them to carry out further investigation into the issue.

This research reminds businesses not to game the review system by heightening the awareness of fictitious entries. They should uphold ethical standards, and refrain from using underhanded tactics of posting fictitious reviews for their offerings, or for those of their competitors. Otherwise, it could lead to what is known as a lemons market that is characterized by information asymmetry between businesses that sell, and users who buy (Akerlof, 1970). Such a market condition will be a lose-lose situation for businesses and users alike.

Given the undeniable existence of foul play, administrators of review websites have their task cut out. Specifically, this research recommends them to pull up their socks in at least four possible ways. First, they could seek help from software companies to explore the possibility to develop systems in order to ascertain review authenticity with reasonable accuracy in real time. Incorporating the findings (cf. Table 4.12, p. 103) in the algorithm could help periodically dust-off reviews that are likely to be fictitious, and bubble those that are likely to be authentic upward on the websites. That said, such an algorithm could ascertain the authenticity of reviews that had already been submitted on the Internet. It might not work for reviews posted at a later time because spammers would eventually become adept in gaming the system.

Second, review website administrators need to guide users on ways to write authentic reviews. Perhaps due to the lack of adequate guidance on ways to write, authentic reviews often end up carrying the traits of fictitious ones as discussed earlier in Section 4.4 (p. 96). Websites should inform users about the ideal levels of comprehensibility, exaggeration, specificity and tentativeness in reviews. They could also encourage users to submit photos and/or videos as a part of reviews. Such multimedia content might be more difficult to manipulate than texts. These guidelines should be disseminated only to bona fide users in order to prevent alerting potential spammers. Alternatively, they could be presented as real-time pop-up messages while users type their reviews. This alternative strategy would, however, require that only bona fide users are given access to the review submission form.

Third, given that prevention is better than cure, review website administrators need to prevent the possibility of submitting fictitious reviews in the first place. By preventing

fictitious reviews from being posted, the need to cure reviews by sifting the grain from the chaff would be obviated. For this purpose, review websites such as TripAdvisor.com should allow submission of reviews only from bona fide users after monetary transactions had been made. Linking reviews to their purchases would at least ensure that the entries had been contributed after post-purchase experiences (Hunt, 2015). This may inspire greater confidence among users in trusting reviews.

Fourth, review website administrators need to guide users on ways to discern review authenticity. This is important because regardless of the extent of filtering done, fictitious reviews could still trickle out occasionally on the Internet. Yet, users hardly receive any guidance on ways to distinguish between authentic and fictitious reviews. Hence, users should minimally be advised to be cautious while reading reviews instead of embracing their truth bias. They could also be guided according to the intervention developed in this research (Appendix G, p. 207). These guidelines could be inserted in the frequently-asked-questions section of review websites. Additionally, they could be used as educational materials in outlets such as the National Consumers League, a well-known organization that helps users avoid online scams (Albert, 2002). They could even be incorporated in websites of national agencies to promote awareness among Internet users (e.g., Government of Canada, 2014). The research contributions to both theory and practice are summarized in Table 7.2.

Table 7.2: Summary of the research contributions.

List of the Research Contributions	
Theoretical Contributions	This research develops the theoretical model of authentic and fictitious reviews (TAF).
	This research amalgamates compartmentalized bodies of knowledge on review authenticity.
	This research conceptualizes authentic and fictitious reviews by widening their scope compared with previous works.
	This research offers a methodological template to develop a ground truth of review authenticity.
	This research presents a holistic operationalization of linguistic cues that help distinguish between authentic and fictitious reviews.
	This research extends the scholarly understanding of the role of guidelines in information processing.
Practical Contributions	This research cautions users not to take the authenticity of reviews for granted.
	This research reminds users to refrain from posting fictitious reviews.
	This research has the potential to identify businesses that manipulate reviews.
	This research reminds businesses not to game the review system.
	This research encourages administrators of review websites to find ways to minimize the chances of users being deceived by fictitious entries.

7.3. Ethical Considerations

Three steps were taken to ensure the ethical soundness of this research. First, all authentic reviews that were collected from the Internet for this research were anonymized. This enabled dissociating the entries from the corresponding reviewers. Only the nationality of the reviewers was noted. Thus, the reviewers who submitted authentic reviews in this research remain unidentifiable. This in turn helps protect their privacy and confidentiality.

Second, all humans who participated in this research did so voluntarily based on goodwill. Vulnerable groups such as children, elderly or those suffering from serious ailments were not involved. Participants were free to discontinue their involvement at any point in time without offering any explanation. No personal identifiers such as names were obtained from the participants to protect their privacy and confidentiality. Approval

was obtained from the Institutional Review Board of the local university prior to studying perceptions of participants in order to ensure strict adherence to ethical standards.

Third, even though this research used hotel reviews as the test case for investigation, no hotel could be unfairly advantaged or disadvantaged. The fictitious reviews created for a given hotel are not meant to be made publicly available by posting them on the Internet either for the same hotel, or for a different hotel. Furthermore, reviews as well as participants' perceptions of reviews were analyzed at an aggregated level, and not at the level of hotels. Thus, this study paints neither an unfairly commendable picture nor an unduly critical image for any property.

7.4. Limitations

The findings of this research should be viewed in light of four limitations. First, it employed a relatively covert approach in collecting authentic reviews than the one used to collect fictitious reviews. As indicated earlier in Section 3.3 (p. 45), authentic reviews were collected from authenticated review websites while fictitious ones were solicited from participants for the creation of the ground truth. In other words, the researcher had greater control in collecting fictitious reviews vis-à-vis authentic ones. In consequence, the intrinsic motivation to write reviews was perhaps less among the participants who wrote fictitious entries vis-à-vis the authentic reviewers.

Second, this research could not control for the medium—computers or mobile devices—used to write authentic and fictitious reviews. It might have been more meaningful to investigate actual and perceived differences between authentic and fictitious reviews when all entries had been contributed from either computers or mobile

devices. Another possibility could have included examining authentic and fictitious reviews that were uniformly distributed in terms of the medium of writing. However, this limitation could not be obviated because once authentic reviews are posted on the Internet, their medium of posting becomes impossible to ascertain. The nature of user-generated content prevents obtaining such precise details.

Third, to collect data, this research failed to leverage on probability sampling, which is ideal to make inferences with confidence. As indicated earlier in Section 3.3.2 (p. 50), fictitious reviews were collected using a combination of snowball sampling and maximum variation sampling. Furthermore, as indicated earlier in Section 5.2 (p. 114), data about perceptions of authentic and fictitious reviews were collected from individuals who were recruited using snowball sampling. Caution is advocated in generalizing the findings of this research.

Fourth, in shedding light on users' perceptions of authentic and fictitious reviews, the only individual difference taken into account in this research is epistemic beliefs. It was chosen because individuals' epistemic beliefs have the potential to shape their ability to decide what is authentic, and what is fictitious in online settings (Hofer, 2004; Kammerer et al., 2013; Qiu et al., 2012; Tu et al., 2008). Nonetheless, this research did not examine the impact of other commonly investigated individual differences such as personality or risk propensity.

7.5. Future Research Directions

This research seeks to propel scholarly inquiry by highlighting three ways to extend the TAF model. One way of extension involves uncovering even more linguistic

cues (e.g., sarcasm) to distinguish between authentic and fictitious reviews as well as even more possible ways for users to discern review authenticity. This is important because scholars and spammers seem to be playing a cat-and-mouse game with one another. No sooner do scholars find a way to distinguish between authentic and fictitious reviews than spammers attempt to devise new—often more devious and ingenious—strategies to circumvent it (Hunt, 2015). Hence, spammers will devise several new strategies to write fictitious reviews down the road. Concurrently, human perceptions of reviews may evolve in unforeseeable ways. The time-sensitive nature of the TAF model makes its current version an inevitable candidate for falsification in the future.

A second way to extend the TAF model involves looking for cues beyond language to examine actual as well as perceived differences between authentic and fictitious reviews. Relying on the language of reviews is certainly not a panacea in an era when reviews with multimedia content are gradually growing in popularity and acceptance (Xu et al., 2015). Unlike purely textual reviews, pictorial and video reviews might not be manipulated easily. This calls for extrapolating the TAF model to cater for multimedia content. Nuances across the medium of review posting—computers or mobile devices—could be taken into account. After all, mobile devices allow capturing photos and videos easily to be posted as reviews. Given that the nature of the problem is continually evolving, finding out the longevity of a theory that explains the phenomenon of review authenticity seems imperative.

A third way to extend the TAF model involves linking human perceptions of review authenticity to measurable outcomes such as intention to adopt a review for making purchase decisions (Casaló et al., 2011), or intention to recommend a review as

helpful for others (Cao et al., 2011). This research demonstrates that humans' perceptions of reviews explain perceived authenticity. Hence, their perceptions may also predict such measurable outcomes. This sets an interesting context to examine if perceived authenticity mediates the relation between humans' perceptions of reviews and their intentions to adopt or recommend the entries.

Apart from possible extensions to the TAF model, this research serves as a springboard to help identify two other directions for future research. One direction includes validating the TAF model. For the purpose of validation, this research relied only on reviews for hotels as the test case for investigation. Hotels specifically represent hedonic aspects of consumption. The extent to which the TAF could be effective remains to be seen when validated with utilitarian products such as cameras and printers as well as products such as backpacks and bike helmets that are associated with both hedonic and utilitarian aspects of consumption (Xu et al., 2015).

Another direction includes studying the problem of authentic and fictitious reviews from a psychology perspective. This research notwithstanding, the psyche and the persona of spammers still remains a mystery. Understanding the psychology of spammers in tandem with that of users who write authentic reviews as well as those who read reviews could enrich the literature in this area.

Overall, this research aspires to whet the appetite of scholars for both theoretical and empirical investigation focusing on authentic and fictitious online reviews. Hopefully, such scholarly efforts would pave the way for a healthier, and a more deception-free cyberspace in the long run.

APPENDIX A IDENTIFIED HOTELS

As indicated in Section 3.2.2 (p. 45), 15 hotels were identified to collect data.

Table A1 lists the hotels across tourist destinations, and hotel categories. Table A2 indicates the volume of reviews attracted by the hotels in Agoda.com, Expedia.com and Hotels.com. Table A3 summarizes their aggregated user-assigned review ratings.

Table A1: List of identified hotels.

Tourist Destinations	Hotel Categories	Hotel Names (acronyms used henceforth)
Bangkok	Luxury	lebua at State Tower Bangkok (lebu)
	Budget	Sawasdee Sukhumvit Bangkok (sukh)
	Mid-range	Nasa Vegas (nasa)
Hong Kong	Luxury	L’Hotel Nina et Convention Centre (nina)
	Budget	USA Hostel Hong Kong (usah)
	Mid-range	Casa Hotel Hong Kong (casa)
Kuala Lumpur	Luxury	Furama Bukit Bintang (fura)
	Budget	D’Oriental Inn Kuala Lumpur (dinn)
	Mid-range	Radius International Kuala Lumpur (radi)
Singapore	Luxury	Mandarin Orchard Singapore (mand)
	Budget	Fragrance Hotel Ruby (ruby)
	Mid-range	Ibis Singapore on Bencoolen (ibis)
Tokyo	Luxury	Hotel Metropolitan Tokyo (metr)
	Budget	Horidome Villa Tokyo (hori)
	Mid-range	Shinjuku Washington Hotel Tokyo (wash)

Table A2: Volume of reviews attracted by the selected hotels.

Tourist Destinations	Hotel Categories	Hotel Names	# Reviews Available in the Selected Websites			
			Agoda.com	Expedia.com	Hotels.com	Total
Bangkok	Luxury	lebu	6,348	903	1,010	8,261
	Budget	sukh	951	9	15	975
	Mid-range	nasa	5,802	142	268	6,212
Hong Kong	Luxury	nina	6,603	155	369	7,127
	Budget	usah	3,074	71	165	3,310
	Mid-range	casa	4,023	80	110	4,213
Kuala Lumpur	Luxury	fura	5,973	134	263	6,370
	Budget	dinn	1,864	79	137	2,080
	Mid-range	radi	5,100	81	121	5,302
Singapore	Luxury	mand	6,678	285	319	7,282
	Budget	ruby	2,300	34	47	2,381
	Mid-range	ibis	4,077	258	302	4,637
Tokyo	Luxury	metr	1,393	237	400	2,030
	Budget	hori	741	107	309	1,157
	Mid-range	wash	1,433	388	472	2,293

Table A3: Aggregated user-assigned review ratings of the selected hotels.

Tourist Destinations	Hotel Categories	Hotel Names	Aggregated Review Ratings ^a		
			Agoda.com	Expedia.com	Hotels.com
Bangkok	Luxury	lebu	8.8	4.6	4.5
	Budget	sukh	7.0	3.2	3.4
	Mid-range	nasa	7.2	3.7	3.5
Hong Kong	Luxury	nina	8.1	4.1	4.2
	Budget	usah	6.5	2.2	2.3
	Mid-range	casa	7.3	3.5	3.4
Kuala Lumpur	Luxury	fura	7.8	3.8	3.8
	Budget	dinn	6.9	3.3	3.1
	Mid-range	radi	6.6	3.0	3.1
Singapore	Luxury	mand	8.2	4.1	4.0
	Budget	ruby	6.7	3.3	3.4
	Mid-range	ibis	7.9	4.2	4.1
Tokyo	Luxury	metr	8.4	4.5	4.2
	Budget	hori	7.9	4.2	4.2
	Mid-range	wash	7.2	3.7	3.9

Note. ^a Agoda.com ratings use a 10-point scale. Expedia.com and Hotels.com ratings use a 5-point scale.

APPENDIX B INSTRUCTIONS FOR FICTITIOUS REVIEWS

As indicated in Section 3.3.2 (p. 52), fictitious reviews were solicited from participants via email instructions. A sample email sent to a participant is as follows:

Hello there!

I am a PhD student under Associate Professor Alton Chua at Nanyang Technological University, Singapore. My research deals with creation and analysis of fake online reviews. If you are in the age group between **21 to 45** years and **USE HOTEL REVIEW WEBSITES** such as TripAdvisor.com, Booking.com or Hotels.com, I would like to invite you to participate in my study. It can be expected to take not more than 20 minutes of your time.

Imagine you work for the marketing department of a hotel. Your boss asks you to write fake reviews in English for a few hotels. The reviews must appear **REALISTIC**. Each review must have a title and a text with **MORE THAN 30 WORDS**. The detailed instructions are appended below. Please ensure that **YOU HAVE NOT STAYED** in the hotels for which you are writing the fake reviews.

Be assured that this study is purely academic in nature. All information will be kept strictly confidential. No hotels will be unfairly advantaged or disadvantaged by your fake reviews. Thank you for your contribution. If you have any questions or concerns, feel free to contact me at snehasis002@ntu.edu.sg.

Regards,
Snehasish Banerjee
PhD Student (Information Studies)

Instructions:

Write **NEGATIVE** fake reviews highlighting **DEMERITS** for the following hotels to disgrace their reputation.

Hotel 1: Mandarin Orchard
(<http://www.meritushotels.com/mice/hotelinformation/mandarin-orchard-singapore>), a 5-star hotel in Singapore.

Hotel 2: Furama Bukit Bintang (<http://www.furama.com/bukitbintang/>), a 5-star hotel in Kuala Lumpur.

Please use the following template to send the fake reviews to snehasis002@ntu.edu.sg:

Mandarin Orchard, a 5-star hotel in Singapore

Review title: <write the title of your negative fake review>

Review text (> 30 words): <write the text of your negative fake review>

Furama Bukit Bintang, a 5-star hotel in Kuala Lumpur

Review title: <write the title of your negative fake review>

Review text (> 30 words): <write the text of your negative fake review>

Gender:

Nationality:

Age group in years (21-25/26-35/36-45):

APPENDIX C INCORRECTLY ANNOTATED REVIEWS

As indicated in Section 3.3.2 (p. 57), the polarity of some reviews were annotated incorrectly. Four such reviews are shown in Table C1. Specifically, it presents two incorrectly annotated authentic reviews: one positive review annotated as moderate, and another moderate review annotated as positive. Besides, it presents two incorrectly annotated fictitious reviews: one moderate review annotated as negative, and another negative review annotated as moderate.

Table C1: Examples of incorrectly annotated reviews.

Ground Truth	Review	Rating	Annotation
Authentic	Title: Great service with convenient location Description: On my first night, I realised that the room wasnt clean. I found some leftovers under the bed. It was really bad but the manager came up and apologised and complimented me free internet for the duration my stay. The next day I found a basket of fruits and chocolates as a token of apology. Upon check out, they apologised again. Overall, GREAT SERVICE!	Positive	Moderate
	Title: Sleek and modern hotel with fantastic views Description: We stayed here for 3 days on out honeymoon and would recommend it to anyone. Our room like the rest of the hotel was sleek and modern and had fantastic views. Our room had been upgraded to a larger one so not sure what size the normal room would have been but certainly ours was fantastic with views our of 2 floor to ceiling windows.	Moderate	Positive
Fictitious	Title: only better than a hostel Description: The room was extremely small and expensive! We felt extremely cramped. I travelled with my husband and we had difficulties opening two suitcases at the same time. May be better than just a hostel. I strongly recommend you look somewhere else to stay. Avoid this place!	Moderate	Negative
	Title: Review of L'Hotel Nina et Convention Center Description: I just recently visited the hotel and found the ambience cold and distanced - probably due to its not at all modern decor. The rooms left me feeling very distant from home as the decoration was very functional. This made my work trip even more stressful even though I tried to hit the less-spacious gym as and when I could.	Negative	Moderate

APPENDIX D DETAILED CLASSIFICATION RESULTS

As indicated in Section 4.2 (p. 80), classification was done using voting among five classification techniques—C4.5 decision tree (C4.5), JRip, logistic regression (LogReg), random forest (RF), and support vector machine (SVM). The detailed classification results for the linguistic framework of cues are presented below. Table D1 presents the volumes of true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP). Table D2 summarizes the performance in terms of six metrics, namely, precision, sensitivity, specificity, accuracy, F1-measure, and area under receiver operating characteristic curve (AUC).

Table D1: Results for true positives, false negatives, true negatives, and false positives.

Classification Techniques	# Accurately Classified	# Inaccurately Classified	# Accurately Classified	# Inaccurately Classified
	Authentic Reviews (TP)	Authentic Reviews (FN)	Fictitious Reviews (TN)	Fictitious Reviews (FP)
LogReg	622	278	668	232
C4.5	625	275	630	270
SVM	612	288	646	254
JRip	625	275	661	239
RF	557	343	712	188
Voting	677	223	714	186

Table D2: Results for performance metrics.

Classification Techniques	Precision	Sensitivity	Specificity	Accuracy	F ₁	AUC
LogReg	0.728	0.691	0.742	71.67%	0.709	0.789
C4.5	0.698	0.694	0.700	69.72%	0.696	0.691
SVM	0.707	0.680	0.718	69.89%	0.693	0.700
JRip	0.723	0.694	0.734	71.44%	0.708	0.747
RF	0.748	0.619	0.791	70.50%	0.677	0.781
Voting	0.784	0.752	0.793	77.28%	0.768	0.853

APPENDIX E SCREENSHOTS OF THE SIMULATED WEBSITE

As indicated in Section 5.1.1 (p. 108), several versions of the simulated review website LoveToTravel.com were designed. Screenshots of three such versions are shown in Figure E1, Figure E2, and Figure E3.

The screenshot shows the LoveToTravel.com website. The header includes a search bar and navigation links: HOME, HOTELS, FLIGHTS, RENTALS, OFFERS, REGISTER, LOGIN, ABOUT US, CONTACTS, and FAQ. The main content area features a section for "Three2One Hotel - a luxury hotel in Hong Kong" with a description: "One of Hong Kong's popular luxury hotel, Three2One offers stylish accommodation and boasts 2 pools and a spa. An extensive range of cuisines is also available. Check-in starts from 14:00, and check-out time is at 12:00." Below this, there are three positive reviews, each with a 5 out of 5 star rating and a title in green:

- 5 out of 5 stars --- "Classy hotel and service with a smile"**
it was a good stay overall. perfect night view, kind hospitality, effective check-in and out, strategic location and nice bar at the top floor. just one small thing for the new receptionist to practice more eye-contact while serving the guests, it will show a little more respect to both parties. credit to the traders personnel serving at the lobby, they are doing great job! they are effective, precise, serve to passion and smile. thank you.
- 5 out of 5 stars --- "Great location with excellent service"**
We had a wonderful stay. We had 4 pple and they were able to us two rooms with a connecting door. The bottom floor had restaurants, bakery and small shops. The staff were very accomodating to our needs. Very clean. Beds very comfortable. Showers excellent.
- 5 out of 5 stars --- "awesome and unbelievable"**
I have stayed in many hotels before. If I have to choose the best hotel I have ever stayed in my life, it has to be this. There is no other hotel in the world like this. Simply amazing and awesome – falling short of words to be honest.

Figure E1: A LoveToTravel.com version showing positive reviews for luxury hotel.

The screenshot shows the LoveToTravel.com website. The header includes a search bar and navigation links: HOME, HOTELS, FLIGHTS, RENTALS, OFFERS, REGISTER, LOGIN, ABOUT US, CONTACTS, and FAQ. The main content area features a section for "Three2One Hotel - a budget hotel in Hong Kong" with a description: "One of Hong Kong's popular budget hotel, Three2One offers no-frills accommodation with basic facilities for the budget-conscious travelers. Cheap eateries are available. Check-in starts from 14:00, and check-out time is at 12:00." Below this, there are three negative reviews, each with a 1 out of 5 star rating and a title in green:

- 1 out of 5 stars --- "Claustrophobic"**
I stayed here with my parents on a tour. The room had no window which gave it a dungeon effect which was suffocating. The breakfast was too plain just with a slice of bread and half cooked egg. The bathroom was very badly maintained with no hot water provided. The television offered only local channels. The stairs and nearby rooms were always noisy which made the night unquieting. Would get a better accommodation for this rate elsewhere.
- 1 out of 5 stars --- "worst - feeling like hell"**
This is the worst thing ever feel like hell hotel so small tiny. The room are smelly and so dirty. Most of the staff are not friendly and the service is not efficient at all. Bathroom got very wet as there was not any shower door.
- 1 out of 5 stars --- "Good value for money but hotel street not ok"**
first look at the actual street of hotel turned us off, we almost transferred to another hotel. it looked like a backstreet of clubs etc. was given a small room with no windows but that's ok as i do not think We will have a good view. room smelled of cigarette smoke. but because its a small hotel, staff could attend to us better and they were accomodating like we bought food outside and ate it at hotel. overall, it was good value for our money but could be better.

Figure E2: A LoveToTravel.com version showing negative reviews for budget hotel.

HOME	HOTELS	FLIGHTS	RENTALS	OFFERS	REGISTER	LOGIN	ABOUT US	CONTACTS	FAQ
<h3>Three2One Hotel - a mid-range hotel in Hong Kong</h3> <p>One of Hong Kong's popular mid-range hotel, Three2One offers highly comfortable accommodation at a reasonable price. A variety of cuisines is also available. Check-in starts from 14:00, and check-out time is at 12:00.</p>									
<div> <p>Posted 1 day ago</p> </div> <div> <p>3 out of 5 stars --- "Quite acceptable"</p> <p>I had a bad first impression with the small room. It feels like staying in school kids' room. It's also smelly along the hallway. But besides that, everything is quite ok. It's very convenient to go sightseeing without the needs to travel long way. One night during my stay it was raining heavily so I had to stay inside the room. Luckily the wifi is really good so I had great time chatting with friends and update my FB with photos. And one very impressive point is that you can go straight to the hotel from airport and vice versa. I'll probably choose it again.</p> </div>									
<div> <p>Posted 1 day ago</p> </div> <div> <p>3 out of 5 stars --- "Unless you have no choice!!!"</p> <p>If you have choices don't choose this hotel as it's not very good. The room I stayed doesn't looks like a hotel room but a cheap hostel room. Inside the room it's quite dark and smelly (not very strong though). Internet is not free and the free wifi in the lobby is slow. Breakfast is provided. It is cheap, but other choices could be better.</p> </div>									
<div> <p>Posted 1 day ago</p> </div> <div> <p>3 out of 5 stars --- "Well kept hotel"</p> <p>the hotel is not new but it is well kept. the only silly thing was there was no shade between the main building annexe building which made walking to the main building a hassle when it is raining. all the amenities were in the main building. they should consider making an underground passage way or a canopy between the main building and annexe building.</p> </div>									

Figure E3: A LoveToTravel.com version showing moderate reviews for mid-range hotel.

APPENDIX F ADVERTISEMENT FOR STUDY PARTICIPATION

As indicated in Section 5.2 (p. 115), an advertisement was used to recruit participants for the survey conducted in the User Study. It is as follows:

Take Part in a Paid Study

We are recruiting students to participate in a study that investigates consumers' willingness to trust online user-generated reviews.

WHO: Students aged between 21 and 45 years having experience of using review websites such as **Amazon** or **TripAdvisor** can participate in this study.

WHAT: This study requires you to complete an online questionnaire. It should not take more than 20 to 30 minutes of your time.

RISK: There are no known risks for participating in this study.

INCENTIVE: SGD10

CONTACT: Mr. Snehasish Banerjee
Phone: +65 8504 4386
Email: Snehasis002@ntu.edu.sg

NOTE: This study is approved by the NTU-Institutional Review Board (IRB code: IRB-2014-10-046). If you have any questions, feel free to contact a representative of the IRB (irb@ntu.edu.sg, or +65 6592 2495).

APPENDIX G GUIDELINE FOR THE ROBUSTNESS CHECK

As indicated in Section 5.5 (p. 139), a guideline was used for the Robustness Check to inform humans how to distinguish between authentic and fictitious reviews. The guideline shown to annotators is presented below, and pictorially depicted in Figure G1.

Step 1: Check if the review is exaggerated with rich use of emotional expressions such as “*awesome*” and “*awful*,” especially negative emotion words such as “*bad*,” firm words such as *always*,” “*never*” and “*perfect*,” function words such as “*as*” and “*are*” as well as punctuations such as exclamation marks.

- If yes, it is fake. Else, go to Step 2.

Step 2: Check if the review—instead of describing specific hotel characteristics—is vague by describing personal experiences with rich use of pronouns such as “*our*,” “*we*,” and “*you*.”

- If yes, it is fake. Else, go to Step 3.

Step 3: Check if the review uses long words such as “*claustrophobic*” and “*disappointing*,” or if it is rich in first person singular words such as “*I*” and “*me*.”

- If yes, it is fake. Else, it is authentic.

Figure G1: Pictorial representation of the guideline as shown to annotators.

APPENDIX H DETAILS OF THE USER DATASET

As indicated in Section 6.1 (p. 150), additional descriptive statistics of the dataset used for the User Study are presented below. Specifically, Table H1 presents the inter-correlations among all the constructs. Table H2 shows the weightings of the items that measured formative constructs while Table H3 shows the loadings of the items that measured reflective constructs. Finally, Table H4 illustrates the cross-loadings.

Table H1: Inter-correlations among constructs.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PCom (1)	1.00									
PExa (2)	-0.16	1.00								
PSpe (3)	0.50	-0.23	1.00							
PTen (4)	-0.14	0.41	-0.19	1.00						
PA (5)	0.29	-0.48	0.46	-0.24	1.00					
EpiR (6)	0.18	-0.04	0.15	-0.02	0.15	1.00				
EpiJ (7)	-0.27	0.03	-0.10	0.05	-0.08	-0.21	1.00			
Anx (8)	-0.18	0.12	-0.08	0.21	-0.05	-0.19	0.35	1.00		
Fam (9)	0.10	0.02	0.06	-0.02	0.03	0.30	-0.31	-0.27	1.00	
Att (10)	0.13	-0.04	0.08	0.03	0.07	0.40	-0.22	-0.16	0.52	1.00

Note. PCom: Perceived comprehensibility, PExa: Perceived exaggeration, PSpe: Perceived specificity, PTen: Perceived tentativeness, PA: Perceived authenticity, EpiR: Epistemic beliefs with respect to perceived reliability of knowledge, EpiJ: Epistemic beliefs with respect to perceived justification for knowing, Anx: Internet anxiety, Fam: Website familiarity, Att: Attitude toward reviews.

Table H2: Item weightings for formative constructs.

Constructs	Questionnaire Items	Weightings	t-Stat
Perceived comprehensibility	PCom1: The title of the review is easy to read.	0.14	0.37
	PCom2: The description of the review is easy to read.	0.92	3.46*
Perceived exaggeration	PExa1: The title of the review is exaggerated.	0.19	0.73
	PExa2: The description of the review is exaggerated.	0.88	4.44*
Perceived specificity	PSpe1: The title of the review is informative.	0.12	0.51
	PSpe2: The description of the review is informative.	0.93	6.01*
Perceived tentativeness	PTen1: The title of the review appears tentative.	0.21	0.37
	PTen2: The description of the review appears tentative.	1.14	2.19*

Note. * represents statistically significant item weighting at the $p < 0.05$ level.

Table H3: Item loadings for reflective constructs.

Constructs	Questionnaire Items	Loadings	t-Stat
Perceived authenticity	PA1: The review is a genuine account of post-trip experience.	0.93	36.30*
	PA2: The review is written after a stay in the hotel.	0.93	32.32*
	PA3: The review is an honest description of a stay in the hotel.	0.94	62.58*
Epistemic beliefs: Reliability	EpiR1: Online reviews on the Internet reflect accurate knowledge about hotels.	0.66	3.72*
	EpiR2: Online reviews provide me with most of the knowledge I need to select a hotel.	0.69	3.58*
	EpiR3: Correct information about hotels can be found in online reviews.	0.79	4.05*
	EpiR4: In online reviews, the richness of detail about hotels is most prominent.	0.61	3.30*
	EpiR5: Online reviews provide correct information about hotels.	0.73	3.63*
	EpiR6: The most important aspect of online reviews is that they contain specific facts about hotels.	0.66	3.48*
	EpiR7: I am most confident that I have selected the appropriate hotel when I have used online reviews as the source of information.	0.60	3.25*
	EpiR8: Most of what is true about hotels are available in online reviews.	0.68	3.45*
	EpiR9: Online reviews contain concrete information about hotels.	0.73	4.13*
	EpiR10: The strength of online reviews is the vast amount of detailed information available about hotels.	0.63	3.39*
	EpiR11: Correct evaluation of hotels is available in online reviews.	0.78	3.98*
Epistemic beliefs: Justification	EpiJ1: To check the credibility of hotel-related information available in online reviews, I try to compare multiple sources. (R)	0.86	3.48*
	EpiJ2: I check if the hotel-related information available in online reviews is logical. (R)	0.82	3.08*
	EpiJ3: To check if the hotel-related information available in online reviews is reliable, I evaluate it in relation to other knowledge I have. (R)	0.80	3.04*
	EpiJ4: I evaluate hotel-related information available in online reviews by checking more sources. (R)	0.78	2.79*
Internet anxiety	Anx1: I hesitate to use the Internet for fear of making mistakes.	0.93	4.28*
	Anx2: I avoid the Internet because it is intimidating to me.	0.94	4.52*
Website familiarity	Fam1: I am familiar with the use of review websites such as Amazon.com and TripAdvisor.com.	0.95	3.31*
	Fam2: I visit review websites such as Amazon.com and TripAdvisor.com.	0.87	3.22*
Attitude toward reviews	Att1: I generally like reading online reviews for hotels.	0.93	5.34*
	Att2: I generally find it interesting to read online reviews for hotels.	0.95	4.76*

Note. * represents statistically significant item loading at the $p < 0.05$ level.
(R) indicates reverse-coded items.

Table H4: Cross-loadings.

	PCom	PExa	PSpe	PTen	PA	EpiR	EpiJ	Anx	Fam	Att
PCom1	0.60	-0.06	0.36	-0.13	0.17	0.19	-0.32	-0.25	0.13	0.11
PCom2	0.99	-0.16	0.49	-0.13	0.28	0.16	-0.24	-0.16	0.08	0.13
PExa1	-0.12	0.69	-0.14	0.24	-0.33	-0.02	0.05	0.09	0.00	-0.04
PExa2	-0.16	0.99	-0.23	0.11	-0.47	-0.04	0.03	0.12	0.02	-0.04
PSpe1	0.34	-0.15	0.61	-0.10	0.28	0.16	-0.07	-0.03	0.05	0.14
PSpe2	0.50	-0.23	0.99	-0.19	0.46	0.14	-0.10	-0.08	0.06	0.07
PTen1	-0.10	0.23	-0.10	0.40	-0.09	0.02	0.05	0.08	-0.02	0.04
PTen2	-0.14	0.41	-0.19	0.98	-0.23	-0.02	0.06	0.20	-0.02	0.04
PA1	0.27	-0.44	0.42	-0.24	0.93	0.15	-0.07	-0.05	0.03	0.08
PA2	0.25	-0.40	0.44	-0.19	0.93	0.12	-0.08	-0.04	0.01	0.04
PA3	0.28	-0.49	0.44	-0.23	0.94	0.16	-0.07	-0.05	0.03	0.08
EpiR1	0.08	0.01	0.01	0.01	0.10	0.66	-0.10	-0.04	0.19	0.23
EpiR2	0.14	-0.03	0.11	0.02	0.12	0.69	-0.17	-0.18	0.20	0.36
EpiR3	0.16	-0.01	0.12	-0.02	0.12	0.79	-0.21	-0.17	0.32	0.32
EpiR4	0.13	-0.02	0.08	-0.05	0.11	0.61	-0.13	-0.18	0.16	0.22
EpiR5	0.07	-0.01	0.09	-0.01	0.10	0.73	-0.08	-0.09	0.23	0.26
EpiR6	0.18	-0.05	0.14	-0.08	0.13	0.66	-0.16	-0.20	0.12	0.30
EpiR7	0.11	-0.02	0.10	0.07	0.07	0.60	-0.09	-0.11	0.24	0.38
EpiR8	0.10	-0.04	0.09	0.00	0.10	0.68	-0.15	-0.06	0.18	0.25
EpiR9	0.07	-0.03	0.08	-0.02	0.11	0.73	-0.10	-0.03	0.20	0.22
EpiR10	0.13	-0.04	0.08	-0.03	0.10	0.63	-0.16	-0.17	0.13	0.20
EpiR11	0.13	-0.04	0.11	-0.04	0.10	0.78	-0.22	-0.19	0.30	0.31
EpiJ1	-0.22	0.01	-0.08	0.05	-0.07	-0.14	0.86	0.30	-0.26	-0.11
EpiJ2	-0.24	0.00	-0.11	0.06	-0.06	-0.18	0.82	0.29	-0.28	-0.18
EpiJ3	-0.18	0.04	-0.05	0.06	-0.06	-0.18	0.80	0.24	-0.28	-0.27
EpiJ4	-0.23	0.04	-0.09	0.02	-0.07	-0.19	0.78	0.30	0.22	-0.16
Anx1	-0.16	0.11	-0.09	0.17	-0.05	-0.16	0.28	0.93	-0.26	-0.13
Anx2	-0.17	0.12	-0.06	0.22	-0.05	-0.19	0.37	0.94	-0.25	-0.16
Fam1	0.13	0.01	0.06	-0.04	0.03	0.29	-0.34	-0.29	0.95	0.46
Fam2	0.03	0.03	0.06	0.01	0.02	0.25	-0.20	-0.18	0.87	0.52
Att1	0.12	-0.04	0.08	0.04	0.06	0.38	-0.22	-0.16	0.54	0.93
Att2	0.13	-0.04	0.08	0.02	0.08	0.37	-0.20	-0.14	0.45	0.95

Note. Shaded cells highlight the correlations between constructs and items measuring those constructs.

APPENDIX I PARTICIPANTS' QUOTATIONS

As indicated in Section 6.4 (p. 162), the qualitative analysis revealed five strategies used by the participants—assessing comprehensibility, checking the level of exaggeration, examining specificity, looking for conflicting information, and relying on intuition. Five sample quotations related to each strategy are presented in Table I1.

Table I1: Quotations about the strategies uncovered using qualitative analysis.

Strategies	Quotations
Comprehensibility	<ul style="list-style-type: none"> • “disorganized write up” (participant 78) • “The statements are . . . not well written” (participant 172) • “very unclear about what the reviewer wants to say” (participant 238) • “sentence structure is bad” (participant 260) • “it stated [its arguments] clearly” (participant 322)
Exaggeration	<ul style="list-style-type: none"> • “more like an advertisement” (participant 151) • “a bit too exaggerated in tone” (participant 163) • “a balanced opinion” (participant 203) • “overly enthusiastic” (participant 266) • “Did not try to understate or overstate” (participant 335)
Specificity	<ul style="list-style-type: none"> • “details...gives the perception of authenticity” (participant 6) • “too little details” (participant 18) • “many details” (participant 28) • “very generic and hardly provide any information” (participant 189) • “quite contextual” (participant 196)
Conflicting information	<ul style="list-style-type: none"> • “[sentences] on hotel services [self-contradictory]” (participant 79) • “[title] matched with the description” (participant 258) • “What was written differs from the rest of the reviews” (participant 312) • “does not even mention about...like the other 2 reviews” (participant 322) • “[arguments do not] match the ratings” (participant 347)
Intuition	<ul style="list-style-type: none"> • “sounds quite believable” (participant 54) • “seems fishy” (participant 168) • “felt authentic” (participant 196) • “seems like a believable story” (participant 227) • “seems not to be a fake one” (participant 331)

REFERENCES

- Abulaish, M., & Bhat, S. Y. (2015). Classifier ensembles using structural features for spammer detection in online social networks. *Foundations of Computing and Decision Sciences*, 40(2), 89-105.
- Afroz, S., Brennan, M., & Greenstadt, R. (2012). Detecting hoaxes, frauds, and deception in writing style online. *IEEE Symposium on Security and Privacy* (pp. 461-475). New York, NY: IEEE.
- Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488-500.
- Albert, M. R. (2002). E-buyer beware: Why online auction fraud should be regulated. *American Business Law Journal*, 39(4), 575-644.
- Alexa. (2013). Top sites by category: Recreation - travel. Retrieved from <http://www.alexa.com/topsites/category/Top/Recreation/Travel>
- Allen, E. (2012, October 9). 'Dear staff. We could do with some positive comments': Hotel boss is caught telling his workers to post fake reviews on TripAdvisor. *MailOnline*. Retrieved from <http://www.dailymail.co.uk/news/article-2214974/Hotel-boss-caught-telling-workers-post-fake-reviews-TripAdvisor.html>
- Al-Nasser, M., Yusoff, R. Z., Islam, R., & AlNasser, A. (2013). Cultural differences identification and its effect on e-service quality perception. *American Journal of Economics and Business Administration*, 5(2), 74-83.
- Ananthakrishnan, S., Vembu, A. N., & Prasad, R. (2011). Model-based parametric features for emotion recognition from speech. *Workshop on Automatic Speech Recognition and Understanding* (pp. 529-534). New York, NY: IEEE.
- Anderson, E. T., & Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, 51(3), 249-269.
- Anderson, J. (1983). LIX and RIX: Variations on a little-known readability index. *Journal of Reading*, 26(6), 490-496.
- Aramberri, J. (2009). The future of tourism and globalization: Some critical remarks. *Futures*, 41(6), 367-376.

- Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior*, 28, 515-526.
- Ascaniis, S. D., & Gretzel, U. (2012). What's in a travel review title? In M. Fuchs, F. Ricci, & L. Cantoni (Eds.), *Information and communication technologies in tourism* (pp. 494-505). Vienna, Austria: Springer-Verlag.
- Au, N., & Akiz, E. H. (2009). Issues and opportunities of Internet hotel marketing in developing countries. *Journal of Travel & Tourism Marketing*, 26(3), 225-243.
- Aune, R. K., Levine, T. R., Park, H. S., Asada, K. J. K., & Banas, J. A. (2005). Tests of theory of communicative responsibility. *Journal of Language and Social Psychology*, 24(4), 358-381.
- Ayello, E. A. (1993). A critique of the AHCPR's 'Preventing pressure ulcers - A patient's guide' as a written instructional tool. *Advances in Skin and Wound Care*, 6(3), 44-51.
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5), 412-424.
- Bambauer-Sachse, S., & Mangold, S. (2011). Brand equity dilution through negative online word-of-mouth communication. *Journal of Retailing and Consumer Services*, 18(1), 38-45.
- Bambauer-Sachse, S., & Mangold, S. (2013). Do consumers still believe what is said in online product reviews? A persuasion knowledge approach. *Journal of Retailing and Consumer Services*, 20(4), 373-381.
- Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to causal modeling: Personal computer adoption and use as an illustration. *Technology Studies*, 2(2), 285-309.
- Bates, M. J. (1979). Information search tactics. *Journal of the American Society for Information Science*, 30(4), 205-214.

- Bates, M. (2005). An introduction to metatheories, theories, and models. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of information behavior* (pp. 1-24). Medford, NJ: Information Today.
- Begeny, J. C., & Greene, D. J. (2014). Can readability formulas be used to successfully gauge difficulty of reading materials? *Psychology in the Schools*, 51(2), 198-215.
- Belkin, N. J. (1990). The cognitive viewpoint in information science. *Journal of Information Science*, 16(1), 11-15.
- Berger, J., & Iyengar, R. (2013). Communication channels and word of mouth: How the medium shapes the message. *Journal of Consumer Research*, 40(3), 567-579.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London, England: Longman.
- Bjering, E., Havro, L.J., & Moen, Ø. (2015). An empirical investigation of self-selection bias and factors influencing review helpfulness. *International Journal of Business and Management*, 10(7), 16-30.
- Bjornsson, C. H. (1983). Readability of newspapers in 11 languages. *Reading Research Quarterly*, 18(4), 480-497.
- Blair, J. P., Levine, T. R., & Shaw, A. S. (2010). Content in context improves deception detection accuracy. *Human Communication Research*, 36(3), 423-442.
- Blanco, C. F., Sarasa, R. G., & Sanclemente, C. O. (2010). Effects of visual and textual information in online product presentations: Looking for the best combination in website design. *European Journal of Information Systems*, 19(6), 668-686.
- Blanke, J., & Chiesa, T. (2011). The travel & tourism competitiveness index 2011: Assessing industry drivers in the wake of the crisis. In J. Blanke, & T. Chiesa (Eds.), *The travel & tourism competitiveness report 2011: Beyond the downturn* (pp. 3-34). Geneva, Switzerland: World Economic Forum.
- Blanzieri, E., & Bryl, A. (2008). A survey of learning-based techniques of email spam filtering. *Artificial Intelligence Review*, 29(1), 63-92.
- Boals, A., & Klein, K. (2005). Word use in emotional narratives about failed romantic relationships and subsequent mental health. *Journal of Language and Social Psychology*, 24(3), 252-268.

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110(2), 305-314.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214-234.
- Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology*, 19(3), 313-329.
- Boush, D.M., & Loken, B. (1991). A process-tracing study of brand extension evaluation. *Journal of Marketing Research*, 28(1), 16-28.
- Box, G., Hunter, W., & Hunter, J. (1978). *Statistics for experimenters*. New York, NY: Wiley.
- Bråten, I., Strømsø, H. I., & Samuelstuen, M. S. (2005). The relationship between Internet-specific epistemological beliefs and learning within Internet technologies. *Journal of Educational Computing Research*, 33(2), 141-171.
- Brookes, B. C. (1968). The derivation and application of the Bradford-Zipf distribution. *Journal of Documentation*, 24(4), 247-265.
- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, 1(1), 8-22.
- Bucci, W., & Freedman, N. (1981). The language of depression. *Bulletin of the Menninger Clinic*, 45(4), 334-358.
- Buelens, H., Totté, N., Deketelaere, A., & Dierickx, K. (2007). Electronic discussion forums in medical ethics education: The impact of didactic guidelines and netiquette. *Medical Education*, 41(7), 711-717.
- Burgoon, J. K., & Qin, T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology*, 25(1), 76-96.
- Burgoon, J. K., Blair, J. P., Qin, T., & Nunamaker, J. F., Jr. (2003). Detecting deception through linguistic analysis. In H. Chen, R. Miranda, D. D. Zeng, C. Demchak, J. Schroeder, & T. Madhusudan (Eds.), *Lecture notes in computer science: Vol.*

2665. *Intelligence and security informatics* (pp. 91-101). Berlin, Germany: Springer-Verlag.
- Burgoon, J., Mayew, W.J., Giboney, J.S., Elkins, A.C., Moffitt, K., Dorn, B., . . . & Spitzley, L. (2016). Which spoken language markers identify deception in high-stakes settings? Evidence from earnings conference calls. *Journal of Language and Social Psychology*, 35(2), 123-157.
- Burla, L., Knierim, B., Barth, J., Liewald, K., Duetz, M., & Abel, T. (2008). From text to codings: Intercoder reliability assessment in qualitative content analysis. *Nursing Research*, 57(2), 113-117.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
- Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach. *Decision Support Systems*, 50(2), 511-521.
- Casaló, L. V., Flavián, C., & Guinalíu, M. (2011). Understanding the intention to follow the advice obtained in an online travel community. *Computers in Human Behavior*, 27(2), 622-633.
- Chall, J. S., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Cambridge, MA: Brookline Books.
- Chan, R. Y. (2004). Consumer responses to environmental advertising in China. *Marketing Intelligence & Planning*, 22(4), 427-437.
- Chang, C. (2016). Responses to conflicting information in computer-mediated communication: Gender difference as an example. *New Media & Society*, 18(1), 5-24.
- Chang, H. H., & Chen, W. S. (2008). The impact of online store environment cues on purchase intention: Trust and perceived risk as a mediator. *Online Information Review*, 32(6), 818-841.
- Chen, J. S. (2001). A case study of Korean outbound travelers’ destination images by using correspondence analysis. *Tourism Management*, 22(4), 345-350.

- Chen, L. S., & Lin, J. Y. (2013). A study on review manipulation classification using decision tree. *Proceedings of the International Conference on Service Systems and Service Management* (pp. 680-685). New York, NY: IEEE.
- Chen, Y. C., Shang, R. A., & Li, M. J. (2014). The effects of perceived relevance of travel blogs' content on the behavioral intention to visit a tourist destination. *Computers in Human Behavior*, 30, 787-799.
- Chen, Z., & Lurie, N. H. (2013). Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research*, 50(4), 463-476.
- Chin, W. W. (1998). Issues and opinion on structural equation modeling. *MIS Quarterly*, 22(1), 7-16.
- Chin, W. W. (2000). Frequently asked questions - Partial least squares & PLS-Graph. Retrieved from <http://disc-nt.cba.uh.edu/chin/plsfaq.htm>
- Chiou, J. S., Hsu, A. C. F., & Hsieh, C. H. (2013). How negative online information affects consumers' brand evaluation: The moderating effects of brand attachment and source credibility. *Online Information Review*, 37(6), 910-926.
- Chiu, Y. L., Liang, J. C., & Tsai, C. C. (2013). Internet-specific epistemic beliefs and self-regulated learning in online academic information searching. *Metacognition and Learning*, 8(3), 235-260.
- Choudhury, V., & Karahanna, E. (2008). The relative advantage of electronic channels: A multidimensional view. *MIS Quarterly*, 32(1), 179-200.
- Cohen, E. (1988). Traditions in the qualitative sociology of tourism. *Annals of Tourism Research*, 15(1), 29-46.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60(2), 283-284.
- Collingridge, D. S., & Gantt, E. E. (2008). The quality of qualitative research. *American Journal of Medical Quality*, 23(5), 389-395.

- Connors, L., Mudambi, S.M., & Schuff, D. (2011). Is it the review or the reviewer? A multi-method approach to determine the antecedents of online review helpfulness. *Hawaii International Conference on System Sciences* (pp. 1-10). New York, NY: IEEE.
- Cook, T. D., & Campbell, D. T. (1976). The design and conduct of quasi-experiments and true experiments in field settings. In M. D. Skokie (Ed.), *Handbook of industrial and organizational psychology* (pp. 223-336). Chicago, IL: Rand McNally.
- Courtis, J. K. (1987). Fry, Smog, Lix and Rix: Insinuations about corporate business communications. *Journal of Business Communication*, 24(2), 19-27.
- Crawford, M., Khoshgoftaar, T. M., Prusa, J. D., Richter, A. N., & Al Najada, H. (2015). Survey of review spam detection using machine learning techniques. *Journal of Big Data*, 2(1), 1-24.
- Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches*. California, CA: Sage.
- Cukier, W., Ngwenyama, O., Bauer, R., & Middleton, C. (2009). A critical analysis of media discourse on information technology: Preliminary results of a proposed method for critical discourse analysis. *Information Systems Journal*, 19(2), 175-196.
- Dawes, J. (2002). Five point vs eleven point scales: Does it make a difference to data characteristics? *Australasian Journal of Market Research*, 10(1), 39-47.
- Dawes, J. (2008). 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*, 51(1). Retrieved from <http://ssrn.com/abstract=2013613>
- Decrop, A. (1999). Triangulation in qualitative tourism research. *Tourism Management*, 20(1), 157-161.
- Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Marketing Science*, 49(10), 1407-1424.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118.

- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51(3), 629-636.
- Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, 17(4), 263-282.
- Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, 38(2), 269-277.
- Du, N., Huang, H., & Li, L. I. (2013). Can online trading survive bad-mouthing? An experimental investigation. *Decision Support Systems*, 56, 419-426.
- Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, 45(4), 1007-1016.
- Ebesu, A. S., & Miller, M. D. (1994). Verbal and nonverbal behaviors as a function of deception type. *Journal of Language and Social Psychology*, 13(4), 418-442.
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2), 155-174.
- Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Journal*, 14(4), 532-550.
- Ekman, P. (1981). Mistakes when deceiving. *Annals of the New York Academy of Sciences*, 364, 269-278.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry: Interpersonal and Biological Processes*, 32(1), 88-106.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body and the face. *Journal of Personality and Social Psychology*, 29(3), 288-298.
- Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, 45(3), 171-212.

- Erman, J., Mahanti, A., Arlitt, M., Cohen, I., & Williamson, C. (2007). Offline/realtime traffic classification using semi-supervised learning. *Performance Evaluation*, 64(9), 1194-1213.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- Feng, S., Xing, L., Gogar, A., & Choi, Y. (2012). Distributional footprints of deceptive product reviews. *Proceedings of the International Conference on Weblogs and Social Media* (pp. 98-105). Palo Alto, CA: AAAI.
- Flesch, R. F. (1951). *How to test readability*. New York, NY: Harper.
- Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research*, 58, 46-64.
- Fletcher, R. J., Havercamp, S. M., Ruedrich, S. L., Benson, B. A., Barnhill, L. J., & Stavrakaki, C. (2009). Clinical usefulness of the diagnostic manual-intellectual disability for mental disorders in persons with intellectual disability: Results from a brief field survey. *Journal of Clinical Psychiatry*, 70(7), 1-8.
- Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 3, 1289-1305.
- Fornell, C., & Bookstein, F. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, 19(4), 440-452.
- Fornell, C., & Larcker, D. (1981). Structural equation models with unobserved variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Freud, S. (1914). *The psychopathology of everyday life*. London, England: Ernest Benn.
- Friedman, H. S., & Tucker, J. S. (1990). Language and deception. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 257-270). New York, NY: Wiley.
- Fuller, C. M., Biros, D. P., & Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications*, 38(7), 8392-8398.

- Fuller, C. M., Biros, D. P., Twitchell, D. P., & Wilson, R. L. (2015). Real-world deception and the impact of severity. *Journal of Computer Information Systems*, 55(2), 59-67.
- Fusilier, D. H., Montes-y-Gómez, M., Rosso, P., & Cabrera, R. G. (2015). Detecting positive and negative deceptive opinions using PU-learning. *Information Processing & Management*, 51(4), 433-443.
- Galante, A., & Thomson, R. I. (2017). The effectiveness of drama as an instructional approach for the development of second language oral fluency, comprehensibility, and accentedness. *TESOL Quarterly*, 51(1), 115-142.
- Gefen, D., Straub, D., & Boudreau, M. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, 4(1), 1-77.
- Gefen, D., Straub, D. W., & Rigdon, E. E. (2011). An update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly*, 35(2), iii-xiv.
- George, J. F., Tilley, P., & Giordano, G. (2014). Sender credibility and deception detection. *Computers in Human Behavior*, 35, 1-11.
- Gerdes, J., Jr., Stringam, B. B., & Brookshire, R. G. (2008). An integrative approach to assess qualitative and quantitative consumer feedback. *Electronic Commerce Research*, 8(4), 217-234.
- Ghiselli, E. E. (1939). All or none versus graded response questionnaires. *Journal of Applied Psychology*, 23(3), 405-415.
- Ghose, A., & 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*, 23(10), 1498-1512.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Hawthorne, NY: Aldine.

- Godin, S. (2005). *All marketers are liars: The power of telling authentic stories in a low trust world*. New York, NY: Penguin.
- Gokhman, S., Hancock, J., Prabhu, P., Ott, M., & Cardie, C. (2012). In search of a gold standard in studies of deception. *Proceedings of the Workshop on Computational Approaches to Deception Detection* (pp. 23-30). Stroudsburg, PA: ACL.
- Government of Canada. (2014, July 28). Don't buy into fake online endorsements: Not all reviews are from legitimate consumers. *Competition Bureau*. Retrieved from <http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03782.html>
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202.
- Grayson, K., & Martinec, R. (2004). Consumer perceptions of iconicity and indexicality and their influence on assessments of authentic market offerings. *Journal of Consumer Research*, 31(2), 296-313.
- Grazioli, S. (2004). Where did they go wrong? An analysis of the failure of knowledgeable internet consumers to detect deception over the internet. *Group Decision and Negotiation*, 13(2), 149-172.
- Grazioli, S., & Jarvenpaa, S. L. (2003). Consumer and business deception on the Internet: Content analysis of documentary evidence. *International Journal of Electronic Commerce*, 7(4), 93-118.
- Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, 30(3), 611-642.
- Gretzel, U., Yoo, K. H., & Purifoy, M. (2007). *Online travel reviews study: Role and impact of online travel reviews*. A&M University, TX: Laboratory for Intelligent Systems in Tourism.
- Gross, M., & Latham, D. (2012). What's skill got to do with it? Information literacy skills and self-views of ability among first-year college students. *Journal of the American Society for Information Science and Technology*, 63(3), 574-583.

- Gu, B., Park, J., & Konana, P. (2012). Research note – The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Information Systems Research*, 23(1), 182-196.
- Gunning, R. (1952). *The technique of clear writing*. New York, NY: McGraw-Hill.
- Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, 6(2), 3-13.
- Gunsch, M. A., Brownlow, S., Haynes, S. E., & Mabe, Z. (2000). Differential linguistic content of various forms of political advertising. *Journal of Broadcasting & Electronic Media*, 44(1), 27-42.
- Gupta, P., & Harris, J. (2010). How e-WOM recommendations influence product consideration and quality of choice: A motivation to process information perspective. *Journal of Business Research*, 63(9), 1041-1049.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11(1), 10-18.
- Hall, W. S., Scholnick, E. K., & Hughes, A. T. (1987). Contextual constraints on usage of cognitive words. *Journal of Psycholinguistic Research*, 16(4), 289-310.
- Hancock, J. T., Curry, L., Goorha, S., & Woodworth, M. (2005). Automated linguistic analysis of deceptive and truthful synchronous computer-mediated communication. *Proceedings of the Hawaii International Conference on System Sciences* (pp. 1-10). New York, NY: IEEE.
- Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45(1), 1-23.
- Harris, C. G. (2012). Detecting deceptive opinion spam using human computation. *Proceedings for the Workshops on Artificial Intelligence* (pp. 87-93). Palo Alto, CA: AAAI.
- Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review*, 19(4), 307-342.

- Hayati, P., Potdar, V., Talevski, A., Firoozeh, N., Sarenche, S., & Yeganeh, E. A. (2010). Definition of spam 2.0: New spamming boom. *Proceedings of the Conference on Digital Ecosystems and Technologies* (pp. 580-584). New York, NY: IEEE.
- Heller, J., & Campbell, W. (1981). A theoretical model of music perception and talent. *Bulletin of the Council for Research in Music Education*, 66/67, 20-24.
- Heydari, A., ali Tavakoli, M., Salim, N., & Heydari, Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications*, 42(7), 3634-3642.
- Highhouse, S. (2009). Designing experiments that generalize. *Organizational Research Methods*, 12(3), 554-566.
- Ho, T. K., Hull, J. J., & Srihari, S. N. (1994). Decision combination in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1), 66-75.
- Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist*, 39(1), 43-55.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88-140.
- Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, 52(3), 674-684.
- Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, 36(6), 929-944.
- Hunt, K. M. (2015). Gaming the system: Fake online reviews v. consumer law. *Computer Law & Security Review*, 31(1), 3-25.
- Internet World Stats. (2013). Internet world users by language: Top 10 languages. Retrieved from <http://www.internetworldstats.com/stats7.htm>
- Ip, C., Lee, H. A., & Law, R. (2012). Profiling the users of travel websites for planning and online experience sharing. *Journal of Hospitality & Tourism Research*, 36(3), 418-426.

- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218.
- Jeong, H. J., & 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*, 25(1), 2-29.
- Jindal, N., & Liu, B. (2008). Opinion spam and analysis. *Proceedings of the International Conference on Web search and Web Data Mining* (pp. 219-230). New York, NY: ACM.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88(1), 67-85.
- Johnson, P. E., Grazioli, S., Jamal, K., & Berryman, R. G. (2001). Detecting deception: Adversarial problem solving in a low base-rate world. *Cognitive Science*, 25(3), 355-392.
- Johnson, S. M., Smith, P., & Tucker, S. (1982). Response format of the job descriptive index: Assessment of reliability and validity by the multitrait-multimethod matrix. *Journal of Applied Psychology*, 67(4), 500-505.
- Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, 29(3), 1193-1203.
- Keil, M., Tan, B. C. Y., Wei, K. K., Saarinen, T., Tuunainen, V., & Wassenaar, A. (2000). Across-cultural study on escalation of commitment behavior in software projects. *MIS Quarterly*, 24(2), 299-325.
- Keller, E., & Berry, J. (2003). *The influentials: One American in ten tells the other nine how to vote, where to eat, and what to buy*. New York, NY: Simon & Schuster.
- Keller, K. L., & Aaker, D. A. (1992). The effects of sequential introduction of brand extensions. *Journal of Marketing Research*, 29(1), 35-60.
- Kerr, D. (2013, April 15). Samsung probed for allegedly bashing rival HTC online. *CNET News*. Retrieved from http://news.cnet.com/8301-1035_3-57579749-94/samsung-probed-for-allegedly-bashing-rival-htc-online/

- Keshtkar, F., & Inkpen, D. (2012). A hierarchical approach to mood classification in blogs. *Natural Language Engineering*, 18(1), 61-81.
- Kim, J. Y., Kioussis, S., & Molleda, J. C. (2015). Use of affect in blog communication: Trust, credibility, and authenticity. *Public Relations Review*, 41(4), 504-507.
- Kim, S. M., Pantel, P., Chklovski, T., & Pennacchiotti, M. (2006). Automatically assessing review helpfulness. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 423-430). Stroudsburg, PA: ACL.
- Kimura, F., & Shridhar, M. (1991). Handwritten numerical recognition based on multiple algorithms. *Pattern Recognition*, 24(10), 969-983.
- Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, Fog count and Flesch reading ease formula) for navy enlisted personnel: Research Branch Report 8-75. Millington, TN: Naval Air Station. Retrieved from www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA006655
- Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. *Proceedings of the Annual Meeting on Association for Computational Linguistics* (pp. 423-430). Stroudsburg, PA: ACL.
- Knapp, M., Hart, R., & Dennis, H. (1974). An exploration of deception as a communication construct. *Human Communication Research*, 1(1), 15-29.
- Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11(3), 205-217.
- Kumar, A. P., & Valsala, P. (2013). Feature selection for high dimensional DNA microarray data using hybrid approaches. *Bioinformation*, 9(16), 824-828.
- Kwon, O., & Sung, Y. (2012). Shifting selves and product reviews: How the effects of product reviews vary depending on the self-views and self-regulatory goals of consumers. *International Journal of Electronic Commerce*, 17(1), 59-82.

- Lahey, C. E., Kernis, M. H., Heppner, W. L., & Lance, C. E. (2008). Individual differences in authenticity and mindfulness as predictors of verbal defensiveness. *Journal of Research in Personality*, 42(1), 230-238.
- Laing, J., Wheeler, F., Reeves, K., & Frost, W. (2014). Assessing the experiential value of heritage assets: A case study of a Chinese heritage precinct, Bendigo, Australia. *Tourism Management*, 40, 180-192.
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46-70.
- Latta, M. M. (2004). Attunement to the creating process in teaching & learning. *Journal of the Canadian Association for Curriculum Studies*, 2(1), 211-226.
- Latta, M. M., & Buck, G. (2008). Enfleshing embodiment: 'Falling into trust' with the body's role in teaching and learning. *Educational Philosophy and Theory*, 40(2), 315-329.
- Lau, R. Y., Liao, S. Y., Kwok, R. C. W., Xu, K., Xia, Y., & Li, Y. (2011). Text mining and probabilistic language modeling for online review spam detection. *ACM Transactions on Management Information Systems*, 2(4), 25:1-30.
- Lee, C. C., Welker, R. B., & Odom, M. D. (2009). Features of computer-mediated, text-based messages that support automatable, linguistics-based indicators for deception detection. *Journal of Information Systems*, 23(1), 19-24.
- Lehman, D. W., Kovács, B., & Carroll, G. R. (2014). Conflicting social codes and organizations: Hygiene and authenticity in consumer evaluations of restaurants. *Management Science*, 60(10), 2602-2617.
- Li, F., Huang, M., Yang, Y., & Zhu, X. (2011). Learning to identify review spam. *Proceedings of the International Joint Conference on Artificial Intelligence* (pp. 2488-2493). Palo Alto, CA: AAAI.
- Li, H., Ye, Q., & Law, R. (2013). Determinants of customer satisfaction in the hotel industry: An application of online review analysis. *Asia Pacific Journal of Tourism Research*, 18(7), 784-802.

- Licht, M. H. (1995). Multiple regression and correlation. In G. Laurence, & Y. R. Paul (Eds.), *Reading and understanding multivariate statistics* (pp. 19-64). Washington, WA: APA.
- Lim, E. P., Nguyen, V. A., Jindal, N., Liu, B., & Lauw, H. W. (2010). Detecting product review spammers using rating behaviors. *Proceedings of the International Conference on Information and Knowledge Management* (pp. 939-948). New York, NY: ACM.
- Lim, Y. S., & Van Der Heide, B. (2015). Evaluating the wisdom of strangers: The perceived credibility of online consumer reviews on Yelp. *Journal of Computer-Mediated Communication*, 20(1), 67-82.
- Limayem, M., Hirt, S. G., & Cheung, C. M. (2007). How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly*, 31(4), 705-737.
- Lohmöller, J. B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg, Germany: Physica-Verlag.
- Lowden, B. G., & Robinson, J. (1999). A statistical approach to rule selection in semantic query optimisation. In Z. W. Raś, & A. Skowron (Eds.), *Foundations of intelligent systems* (pp. 330-339). Berlin, Germany: Springer-Verlag.
- Lu, Y., Zhang, L., Xiao, Y., & Li, Y. (2013). Simultaneously detecting fake reviews and review spammers using factor graph model. *Proceedings of the Annual Web Science Conference* (pp. 225-233). New York, NY: ACM.
- Ludwig, S., de Ruyter, K., Friedman, M., Brüggem, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1), 87-103.
- Luo, J., Ba, S., & Zhang, H. (2012). The effectiveness of online shopping characteristics and well-designed websites on satisfaction. *MIS Quarterly*, 36(4), 1131-1144.
- MacKenzie, I. S., & Zhang, S. X. (1999). The design and evaluation of a high-performance soft keyboard. *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 25-31). New York, NY: ACM.

- Mason, L., & Ariasi, N. (2010). Critical thinking about biology during web page reading: Tracking students' evaluation of sources and information through eye fixations. In L. Verschaffel, E. de Corte, T. de Jong, & J. Elen (Eds.), *Use of external representations in reasoning and problem solving: Analysis and improvement* (pp. 55-73). London, England: Routledge.
- Massung, S., Zhai, C., & Hockenmaier, J. (2013). Structural parse tree features for text representation. *Proceedings of the International Conference on Semantic Computing* (pp. 9-16). New York, NY: IEEE.
- Maurer, C., & Schaich, S. (2011). Online customer reviews used as complaint management tool. In R. Law, M. Fuchs, & F. Ricci (Eds.), *Information and communication technologies in tourism* (pp. 499-511). Vienna, Austria: Springer-Verlag.
- McColl, M. A., & Pranger, T. (1994). Theory and practice in the occupational therapy guidelines for client-centred practice. *Canadian Journal of Occupational Therapy*, 61(5), 250-259.
- McCornack, S. A. (1992). Information manipulation theory. *Communications Monographs*, 59(1), 1-16.
- McEnroe, J. E., & Martens, S. C. (2001). Auditors' and investors' perceptions of the "expectation gap". *Accounting Horizons*, 15(4), 345-358.
- McKnight, H., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems*, 11(3), 297-323.
- McKnight, H., & Kacmar, C. (2006). Factors of information credibility for an internet advice site. *Proceedings of the Annual Hawaii International Conference on System Sciences* (pp. 113b: 1-10). New York, NY: IEEE.
- McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading*, 12(8), 639-646.
- McLeod, J. M., & Pan, Z. (2005). Concept explication and theory construction. In S. Dunwoody, L. Becker, D. McLeod, & G. Kosicki (Eds.), *The evolution of key*

- mass communication concepts: Honoring Jack M. McLeod* (pp. 13-76). Cresskill, NJ: Hampton Press.
- McQuaid, S. M., Woodworth, M., Hutton, E. L., Porter, S., & ten Brinke, L. (2015). Automated insights: Verbal cues to deception in real-life high-stakes lies. *Psychology, Crime & Law*, 21(7), 617-631.
- Mehrabian, A. (1967). Attitudes inferred from non-immediacy of verbal communications. *Journal of Verbal Learning and Verbal Behavior*, 6(2), 294-295.
- Miao, Q., Li, Q., & Zeng, D. (2010). Fine-grained opinion mining by integrating multiple review sources. *Journal of the American Society for Information Science and Technology*, 61(11), 2288-2299.
- Mihalcea, R., & Burzo, M. (2012). Towards multimodal deception detection - Step 1: Building a collection of deceptive videos. *Proceedings of the International Conference on Multimodal Interaction* (pp. 189-192). New York, NY: ACM.
- Mihalcea, R., & Strapparava, C. (2009). The lie detector: Explorations in the automatic recognition of deceptive language. *Proceedings of the International Joint Conference on Natural Language Processing* (pp. 309-312). Stroudsburg, PA: ACL.
- Misra, P., Agarwal, N., Kasabwala, K., Hansberry, D. R., Setzen, M., & Eloy, J. A. (2013). Readability analysis of healthcare-oriented education resources from the American academy of facial plastic and reconstructive surgery. *The Laryngoscope*, 123(1), 90-96.
- Missen, M. M. S., & Boughanem, M. (2009). Using wordnet's semantic relations for opinion detection in blogs. In M. Boughanem, C. Berrut, J. Mothe, & C. Soule-Dupuy (Eds.), *Advances in information retrieval* (pp. 729-733). Berlin, Germany: Springer-Verlag.
- Molina-Azorin, J. F. (2016). Mixed methods research: An opportunity to improve our studies and our research skills. *European Journal of Management and Business Economics*, 25(2), 37-38.

- Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3), 192-222.
- Morris, P. H., & Lewis, D. (2010). Tackling diving: The perception of deceptive intentions in association football (soccer). *Journal of Nonverbal Behavior*, 34(1), 1-13.
- Munzel, A. (2015). Malicious practice of fake reviews: Experimental insight into the potential of contextual indicators in assisting consumers to detect deceptive opinion spam. *Recherche et Applications en Marketing* (English Edition), 30(4), 24-50.
- Nakamura, J. (1991). The relationship among genres in the LOB corpus based upon the distribution of grammatical tags. *JACET Bulletin*, 22, 44-74.
- Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5), 665-675.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690.
- O'Mahony, M. P., & Smyth, B. (2009). Learning to recommend helpful hotel reviews. *Proceedings of the Conference on Recommender systems* (pp. 305-308). New York, NY: ACM.
- Ong, T., Mannino, M., & Gregg, D. (2014). Linguistic characteristics of shill reviews. *Electronic Commerce Research and Applications*, 13(2), 69-78.
- Ott, M., Cardie, C., & Hancock, J. (2012). Estimating the prevalence of deception in online review communities. *Proceedings of the International Conference on World Wide Web* (pp. 201-210). New York, NY: ACM.
- Ott, M., Cardie, C., & Hancock, J. T. (2013). Negative deceptive opinion spam. *North American Chapter of the Association for Computational Linguistics* (pp. 497-501). Stroudsburg, PA: ACL.

- Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. *Proceedings of the Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 309-319). Stroudsburg, PA: ACL.
- Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management*, 32(2), 215-224.
- Park, D., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, 11(4), 125-148.
- Parker, C. (2013). On measuring the performance of binary classifiers. *Knowledge and Information Systems*, 35(1), 131-152.
- Pasupathi, M. (2007). Telling and the remembered self: Linguistic differences in memories for previously disclosed and previously undisclosed events. *Memory*, 15(3), 258-270.
- Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7(3), 69-103.
- Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8(3), 162-166.
- Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science*, 10(3), 90-93.
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count: LIWC [Computer software]. Austin, TX: liwc.net.
- Pérez-Sabater, C., Peña-Martínez, G., Turney, E., & Montero-Fleta, B. (2008). A spoken genre gets written: Online football commentaries in English, French, and Spanish. *Written Communication*, 25(2), 235-261.
- Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly*, 31(4), 623-656.

- Pew Research Centre. (2015, October 8). Social media usage: 2005-2015. Retrieved from <http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/>
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12(4), 531-544.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program: Electronic Library and Information Systems*, 14(3), 130-137.
- Porter, S., & ten Brinke, L. (2008). Reading between the lies identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19(5), 508-514.
- Porter, S., & ten Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology*, 15(1), 57-75.
- Porter, S., ten Brinke, L., & Wallace, B. (2012). Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity. *Journal of Nonverbal Behavior*, 36(1), 23-37.
- Qiu, L., Pang, J., & Lim, K. H. (2012). Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence. *Decision Support Systems*, 54(1), 631-643.
- Ranganathan, C., & Sethi, V. (2002). Rationality in strategic information technology decisions: The impact of shared domain knowledge and IT unit structure. *Decision Sciences*, 33(1), 59-86.
- Ratchford, B. T., Lee, M. S., & Talukdar, D. (2003). The impact of the Internet on information search for automobiles. *Journal of Marketing Research*, 40(2), 193-209.
- Ravid, D. (2005). Emergence of linguistic complexity in later language development: Evidence from expository text construction. In D. Ravid & H. Shyldkrot (Eds.), *Perspectives on language and language development* (pp. 337-355). Springer US.

- Rayson, P., Wilson, A., & Leech, G. (2001). Grammatical word class variation within the British National Corpus sampler. *Language and Computers*, 36(1), 295-306.
- Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, 49(4), 243-256.
- Ridgon, R. E. (1998). Structural equation modelling. In G. A. Marcoulides (Ed.), *Modern methods for business research*. Mahwah, NJ: Lawrence Erlbaum.
- Ridzuan, F., Potdar, V., & Talevski, A. (2010). Key parameters in identifying cost of spam 2.0. *Proceedings of the International Conference on Advanced Information Networking and Applications* (pp. 789-796). New York, NY: IEEE.
- Ringle, C. M., Wende, S., & Will, A. (2005). SmartPLS - Version 2.0 [Computer software]. University of Hamburg: smartpls.de.
- Rodríguez, N. G., Pérez, M. J. S., Gutiérrez, J. A. T. (2008). Can a good organizational climate compensate for a lack of top management commitment to new product development? *Journal of Business Research*, 61(2), 118-131.
- Rong, J., Vu, H. Q., Law, R., & Li, G. (2012). A behavioral analysis of web sharers and browsers in Hong Kong using targeted association rule mining. *Tourism Management*, 33(4), 731-740.
- Rubin, V. L., & Lukoianova, T. (2015). Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology*, 66(5), 905-917.
- Rude, S., Gortner, E. M., & Pennebaker, J. (2004). Language use of depressed and depression vulnerable college students. *Cognition & Emotion*, 18(8), 1121-1133.
- Runfola, A., Rosati, M., & Guercini, S. (2012). New business models in online hotel distribution: Emerging private sales versus leading IDS. *Service Business*, 7(2), 183-205.
- Saldaña, J. (2015). *The coding manual for qualitative researchers*. London, England: Sage.
- Sanderson, J. (2013). From loving the hero to despising the villain: Sports fans, Facebook, and social identity threats. *Mass Communication and Society*, 16(4), 487-509.

- Savolainen, R. (2011). Judging the quality and credibility of information in Internet discussion forums. *Journal of the American Society for Information Science and Technology*, 62(7), 1243-1256.
- Schiffman, H. R. (2001). *Sensation and perception* (5th ed.). New York, NY: Wiley.
- Schrank, B., Sibitz, I., Unger, A., & Amering, M. (2010). How patients with schizophrenia use the internet: Qualitative study. *Journal of Medical Internet Research*, 12(5). Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3057320/>
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, 21(4), 76-94.
- Shapiro, D. L., Buttner, E. H., & Barry, B. (1994). Explanations: What factors enhance their perceived adequacy? *Organizational Behavior and Human Decision Processes*, 58(3), 346-368.
- Shie, J. S. (2010). Lexical feature variations between New York Times and Times Supplement News headlines. *Concentric: Studies in Linguistics*, 36(1), 79-103.
- Shojaee, S., Murad, M.A.A., Azman, A.B., Sharef, N.M., & Nadali, S. (2013). Detecting deceptive reviews using lexical and syntactic features. *Proceedings of the International Conference on Intelligent Systems Design and Applications* (pp. 53-58). New York, NY: IEEE.
- Sidali, K. L., Schulze, H., & Spiller, A. (2009). The impact of online reviews on the choice of holiday accommodations. In W. Höpken, U. Gretzel, & R. Law (Eds.), *Information and communication technologies in tourism* (pp. 87-98). Vienna, Austria: Springer-Verlag.
- Singh, A. (1997). Asia Pacific tourism industry: Current trends and future outlook. *Asia Pacific Journal of Tourism Research*, 2(1), 89-99.
- Small, H. (1999). A passage through science: Crossing disciplinary boundaries. *Library Trends*, 48(1), 72-108.
- Smith, E. A., & Senter, R. J. (1967). *Automated Readability Index*. AMRL-TR-66-22 Wright-Patterson AFB, OH: Aerospace Medical Division.

- Smith, K. W. (1974). On estimating the reliability of composite indexes through factor analysis. *Sociological Methods & Research*, 2(4), 485-510.
- Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6), 1310-1323.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford, England: Blackwell.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359-393.
- Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, 20(4), 421-446.
- Steenkamp, J. B., & Geyskens, I. (2006). How country characteristics affect the perceived value of a website. *Journal of Marketing*, 70(3), 136-150.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63(4), 517-522.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
- Stringam, B. B., & Gerdes, J., Jr. (2010). An analysis of word-of-mouth ratings and guest comments of online hotel distribution sites. *Journal of Hospitality Marketing & Management*, 19(7), 773-796.
- Svider, P. F., Agarwal, N., Choudhry, O. J., Hajart, A. F., Baredes, S., Liu, J. K., & Eloy, J. A. (2013). Readability assessment of online patient education materials from academic otolaryngology - head and neck surgery departments. *American Journal of Otolaryngology*, 34(1), 31-35.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Tenenhaus, M., Amato, S., & Vinzi, V. E. (2004). A global goodness-of-fit index for PLS structural equation modelling. *Proceedings of the XLII SIS Scientific Meeting* (pp. 739-742). Retrieved from <http://old.sis-statistica.org/files/pdf/atti/RSBa2004p739-742.pdf>

- Tharenou, P., Donohue, R., & Cooper, B. (2007). *Management research methods*. Cambridge University Press.
- The Guardian. (2011, January 5). Amazon withdraws ebook explaining how to manipulate its sales rankings. Retrieved from <http://www.guardian.co.uk/books/2011/jan/05/amazon-ebook-manipulate-kindle-rankings>
- Toma, C. L., & D'Angelo, J. D. (2015). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of Language and Social Psychology*, 34(1), 25-45.
- Toma, C. L., & Hancock, J. T. (2012). What lies beneath: The linguistic traces of deception in online dating profiles. *Journal of Communication*, 62(1), 78-97.
- Tsang, A. S., & Prendergast, G. (2009). Is a “star” worth a thousand words? The interplay between product-review texts and rating valences. *European Journal of Marketing*, 43(11/12), 1269-1280.
- Tse, T. S. M. (2013). The marketing role of the internet in launching a hotel: The case of Hotel ICON. *Journal of Hospitality Marketing & Management*, 22(8), 895-908.
- Tsur, O., Davidov, D., & Rappoport, A. (2010). A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. *Proceedings of the International Conference on Weblogs and Social Media* (pp. 162-169). Palo Alto, CA: AAAI.
- Tu, Y. W., Shih, M., & Tsai, C. C. (2008). Eighth graders' web searching strategies and outcomes: The role of task types, web experiences and epistemological beliefs. *Computers & Education*, 51(3), 1142-1153.
- Undeutsch, U. (1967). *Forensische psychologie* [Forensic psychology]. Gottingen, Germany: Verlag fur Psychologie.
- United Nations World Tourism Organization. (2013). International tourism to continue robust growth in 2013. Retrieved from <http://media.unwto.org/en/press-release/2013-01-28/international-tourism-continue-robust-growth-2013>

- Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *JITTA: Journal of Information Technology Theory and Application*, 11(2), 5-39.
- Vartapetian, A., & Gillam, L. (2012). 'I don't know where he is not': Does deception research yet offer a basis for deception detectives? *Workshop on Computational Approaches to Deception Detection* (pp. 5-14). Stroudsburg, PA: ACL.
- Vega, C. (2012, November 19). Yelp outs companies that pay for positive reviews. *abc News*. Retrieved from <http://abcnews.go.com/blogs/business/2012/11/yelp-outs-companies-that-pay-for-positive-reviews/>
- Velayutham, S., Aldridge, J. M., & Fraser, B. (2012). Gender differences in student motivation and self-regulation in science learning: A multi-group structural equation modeling analysis. *International Journal of Science and Mathematics Education*, 10(6), 1347-1368.
- Verhoef, P. C., Franses, P. H., & Hoekstra, J. C. (2002). The effect of relational constructs on customer referrals and number of services purchased from a multiservice provider: does age of relationship matter? *Journal of the Academy of Marketing Science*, 30(3), 202-216.
- Vilares, M. J., Almeida, M. H., & Coelho, P. S. (2010). Comparison of likelihood and PLS estimators for structural equation modeling: A simulation with customer satisfaction data. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 289-305). Berlin, Germany: Springer-Verlag.
- Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51(3), 576-586.
- Viswanathan, M. (2005). *Measurement error and research design*. Thousand Oaks, CA: Sage.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Hoboken, NJ: Wiley.

- Vrij, A., & Baxter, M. (1999). Accuracy and confidence in detecting truths and lies in elaborations and denials: Truth bias, lie bias and individual differences. *Expert Evidence*, 7(1), 25-36.
- Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, 24(4), 239-263.
- Wang, T., & Zhu, H. (2014). Voting for deceptive opinion spam detection. Retrieved from <http://arxiv.org/pdf/1409.4504.pdf>
- Watson, C. (2014). An exploratory study of secondary students' judgments of the relevance and reliability of information. *Journal of the Association for Information Science and Technology*, 65(7), 1385-1408.
- Watson, K. W., & Ragsdale, J. D. (1981). Linguistic indices of truthful and deceptive responses to employment interview questions. *Journal of Applied Communication Research*, 9(2), 59-71.
- Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(3), 236-247.
- Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement*, 34(1), 25-33.
- Whitmire, E. (2004). The relationship between undergraduates' epistemological beliefs, reflective judgment, and their information-seeking behaviour. *Information Processing and Management*, 40(1), 97-111.
- Whitworth, B., & Whitworth, E. (2004). Spam and the social-technical gap. *Computer*, 37(10), 38-45.
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal*, 46(4), 1060-1106.
- Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly*, 25(1), 17-41.

- Wold, H. (1985). Partial least squares. In S. Kotz, & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 6, pp. 581-591), New York, NY: Wiley.
- Wold, H., & Jöreskog, K. G. (1982). *Systems under indirect observation: Causality, structure, prediction* (Vol. 2). Amsterdam, The Netherlands: North-Holland Publishing Company.
- Wu, G., Greene, D., & Cunningham, P. (2010). Merging multiple criteria to identify suspicious reviews. *Proceedings of the ACM Conference on Recommender Systems* (pp. 241-244). New York, NY: ACM.
- Xia, L., & Bechwati, N. N. (2008). Word of mouse: The role of cognitive personalization in online consumer reviews. *Journal of Interactive Advertising*, 9(1), 3-13.
- Xiao, B., & Benbasat, I. (2011). Product-related deception in e-commerce: A theoretical perspective. *MIS Quarterly*, 35(1), 169-196.
- Xu, P., Chen, L., & Santhanam, R. (2015). Will video be the next generation of e-commerce product reviews? Presentation format and the role of product type. *Decision Support Systems*, 73, 85-96.
- Yang, Y., & Pedersen, J. O. (1997). A comparative study on feature selection in text categorization. *Proceedings of the International Conference on Machine Learning* (pp. 412-420). Palo Alto, CA: AAAI.
- Yoo, K. H., & Gretzel, U. (2009). Comparison of deceptive and truthful travel reviews. In W. Höpken, U. Gretzel, & R. Law (Eds.), *Information and communication technologies in tourism* (pp. 37-47). Vienna, Austria: Springer-Verlag.
- Zakaluk, B.L., & Samuels, S.J. (1988). *Readability: Its past, present and future*. Newark, DE: International Reading Association.
- Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 10, 27-36.
- Zhang, K. Z. K., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the moderating effect of inconsistent reviews and its gender differences on consumers' online shopping decision. *International Journal of Information Management*, 34(2), 89-98.

- Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the American Society for Information Science and Technology*, 57(3), 378-393.
- Zhou, L., Burgoon, J. K., & Twitchell, D. P. (2003). A longitudinal analysis of language behavior of deception in e-mail. *Conference on Intelligence and Security Informatics* (pp. 102-110). Berlin, Germany: Springer-Verlag.
- Zhou, L., Shi, Y., & Zhang, D. (2008). A statistical language modeling approach to online deception detection. *IEEE Transactions on Knowledge and Data Engineering*, 20(8), 1077-1081.
- Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation*, 13(1), 81-106.
- Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T., & Nunamaker, J. F., Jr. (2004). A comparison of classification methods for predicting deception in computer-mediated communication. *Journal of Management Information Systems*, 20(4), 139-165.
- Zikmund, W. G. (2003). *Business research methods* (7th ed.). Ohio, OH: Thomson South-Western.