TWO STUDIES ON INDIVIDUALS' PERCEPTION ON ONLINE DISINFORMATION

XU YUKUAN

PhD

The Hong Kong Polytechnic University

This programme is jointly offered by The Hong Kong Polytechnic University and Harbin Institute of Technology

2021

The Hong Kong Polytechnic University

Department of Management and Marketing

Harbin Institute of Technology

School of Economics and Management

Two Studies on Individuals’ Perception on Online Disinformation

XU Yukuan

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

July, 2020

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signed ____________

XU Yukuan

Abstract

Online disinformation has become a relatively common phenomenon on several platforms including social media platforms, ecommerce platforms and news platforms, and causes several serious consequences. In this paper, I explore two of its primary damaging manifestations, namely fake reviews of goods or services, which can mislead customers and potentially diminish firms' long-term profits, and fake news, which are intentionally and verifiably false to mislead readers.

In the first study, I focus on the effect of negative rating deviation on perceived review manipulation and explore the boundary conditions. Existing studies mainly focus on the detection and impact of fake reviews but do not investigate customers' perception of review legitimacy. By introducing a new concept— perceived review manipulation—which evaluates the extent to which a customer regards the opinions or recommendations of a review to be misleading or manipulated, this study investigates the effect of reviews that deviate from average ratings on perceived review manipulation and explores the moderating effect of review content concreteness and reviewer rating distribution. This study also examines whether perceived review manipulation functions as a mechanism that mediates the effect of deviant ratings on perceived review helpfulness/adoption. By conducting two online randomized experiments and one field study on Yelp, the findings suggest: (1) reviews with deviant ratings are more likely to be perceived (filtered) as review manipulation; (2) customers (platforms) are more likely to perceive (classify) reviews with deviant ratings as review manipulations when review content is abstract rather than concrete; (3) customers are more (less) likely to perceive reviews with deviant ratings as review manipulation when reviewer rating distribution is negative (positive); (4) platforms are more (less) likely to filter reviews with deviant ratings as review manipulation when the skewness of reviewer rating distribution is large (small); (5) perceived review manipulation mediates the effect of reviews with deviant ratings on perceived review helpfulness/adoption.

In the second study, I intend to investigate individuals' ability to distinguish fake news from real news in normal situations and in the situations under which different arousal level and personal involvement exist. Fake news has penetrated to individuals' life especially when a particular epidemic event occurs. The reason is that the occurrence of some epidemic events is more likely to motivate individuals to evaluate news based on their sentiments rather than rationality, thus likely lowering individuals' ability to distinguish between real news and fake news. This study focuses on the role of two dimensions of sentiments—arousal level and personal involvement—on individuals' judgement on real news and fake news. By conducting two online randomized experiments in the context of COVID-19, this study finds that individuals have the ability to distinguish fake news and real news in normal situations, while they are more likely to trust or share fake news when news content can trigger high arousal level. For individuals with a high involvement toward COVID-19 news, their trusting perception toward fake news whose content triggering high arousal level becomes even higher.

The findings have implications for academics and practitioners. Theoretically, the findings contribute to online review literature by proposing the new concept of perceived review manipulation and identifying the factors that can influence perceived review manipulation as well as fake news literature by introducing arousal level and personal involvement to the new context and better understanding individuals' evaluations on fake news. Practically, the findings provide new insights for ecommerce platforms and news platforms about the emphasis on online disinformation regulation.

Keywords: online disinformation, review manipulation, deviant ratings, content concreteness, rating distribution, online experiments, field study, fake news, arousal level, personal involvement

Acknowledgements

I would like to take this opportunity to thank all those people who helped me during my Ph.D. journey. Especially, I must express my thanks to my supervisors, Prof. XU Xin and Prof. ZHANG Ziqiong, for their patient guidance and unconditional support to my every project. Without their help, I cannot finish the thesis. I also must express my thanks to BOE members for their valuable suggestions for improving my thesis. My appreciation also goes to the colleagues for their help.

Table of Contents

Lists of Figures

List of Tables

Chapter 1
Introduction

1.1 Research motivation

Online disinformation is defined as “false information that is purposely spread to deceive people.” (Lazer et al., 2018). It is regarded as a major problem in the dissemination of information in many important societal realms, including politics, medicine, and commerce (Waszak et al. 2018; Jang & Kim, 2018; Barbado et al. 2019). The most famous case happens in the context of politics. That is, during the period of 2016 US presidential election, election results were purportedly swayed by fake news on Facebook (Allcott & Gentzkow, 2017). The serious consequences caused by online disinformation make it very essential for scholars to conduct related studies.

Online disinformation has several manifestations. In the ecommerce context, one of the most important forms of online disinformation is fake reviews—i.e., reviews that are not written by real customers or do not present true opinions about product or service quality. Several review websites, such as Yelp, Amazon, etc., suffer from problems generated by fake reviews, which often lead customers to make decisions based on faulty information that can damage the reputation of the platform or firm associated with the reviews (Luca & Zervas, 2016). In other contexts including politics and health areas, fake news is the main form of online disinformation, which mimics the look and feel of news and mislead readers to trust wrong information (Gelfert, 2018). Several platforms including social media platforms such as Facebook and news platforms such as Toutiao.com suffer from fake news, causing the damage of company reputation and even the instability of society (Lazer et al., 2018).

Existing studies on online disinformation mainly falls into several lines, such as its influence (Bentzen, 2019), its reach (Fletcher et al., 2018), its detection (Kumar et al., 2018) and individuals’ resilience (Humprecht et al., 2020). Although advanced technology such as hierarchical supervised learning has been explored to detect online disinformation (Kumar et al., 2018), it can still find a way to mislead individuals as it can be created by everyone and reaches considerable number of users before being filtered. Thus, individuals have to discern online disinformation by themselves. In other words, whether a platform can efficiently control the negative consequences of online disinformation largely depends on its individuals' perception on this kind of information (such as fake review and fake news mentioned above). However, to my best knowledge, studies that focus on determinants of individuals' perception on online disinformation are largely limited. As fake review and fake news are two important representations of online disinformation, I focus on how individuals perceive fake review and the factors can influence individual perception. Going beyond simple individual perception, I also concentrate on the determinants of individuals' ability to judge fake news and real news. Specifically, in this study, I first intend to explore individuals' perceptions on fake reviews with the existence of negative rating deviation and examine boundary conditions under this influence. I also investigate individuals' ability to discern fake news and real news by considering the moderating impact of individuals' emotional aspects.

1.2 Research questions

As the goal of this thesis is to investigate the factors that can influence individuals' perception on online disinformation in different contexts including ecommerce platform and news platform, I summarize the following research questions that will be answered by the next two studies:

RQ1: How individuals perceive reviews with negative rating deviation as being manipulated?

RQ2: How review content impacts individuals' manipulation perception on reviews with negative rating deviation?

RQ3: How reviewer rating distribution influences individuals' manipulation perception on reviews with negative rating deviation?

RQ4: How news content that can trigger a high arousal level influences individuals' ability to distinguish fake news from real news?

RQ5: How individuals' involvement toward news influence individuals' abilities to distinguish fake news from real news?

To solve these research questions, I collect experimental data by recruiting participants from United States and Mainland China and collect objective data from an existing study. Specifically, in Study one, I employ mixed-method by combining an online randomized experiment and a field study. In Study two, I conduct two online randomized experiment as pilot study and formal study.

Chapter 2

What influences customers' manipulation perception of reviews? A mixed-method exploration based on online randomized experiments and a field study

2.1 Introduction

Online reviews are an extremely important tool that customers use to make purchase decisions and are thus highly influential for product sales (Hu et al., 2014; King et al., 2014; Ghose & Ipeirotis, 2012; Chevalier & Mayzlin, 2006). Therefore, practitioners actively seek to encourage customers to share their experience through reviews, resulting in the proliferation of online reviews. However, the prevalent use of online reviews also has a dark side, and reviews that are low quality or misleading can have detrimental effects (Zhang et al., 2016). Although several studies have already focused on the factors/designs that can help readers evaluate/perceive fake news (Kim & Dennis, 2019; Kim et al., 2019), previous studies that have focused on fake reviews have primarily investigated antecedents, consequences, and interventions (Wu et al., 2020), while largely ignoring the perspective of review readers' perceptions of fake reviews. This gap in the literature, combined with the increasing prevalence of fake reviews, motivates the core element of this study: the proposition of a new concept, *perceived review manipulation*, which addresses the extent to which a customer regards opinions or recommendations in a review as misleading or manipulated.

To help consumers understand peers' evaluations more quickly and easily, given a large amount of information, platforms provide average ratings of all prior reviews. In some cases, individual ratings are relatively or even extremely negative compared to the average rating (see Figure 1). This deviation may influence consumers' perceptions of individual reviews and, in turn, impact customers' attitudes and purchase behaviors, making it thus essential to explore the effect of rating deviation (Chevalier & Mayzlin, 2006). Some studies found that consistency between a target review and other reviews improves review adoption (Qiu et al., 2012; Cheung et al., 2009), while others suggest that individual reviews with deviant ratings attract more customer attention and are thus perceived as more helpful (Gao et al., 2017; Shen et al., 2015). The presence of competing findings invites further study. I seek to understand whether reviews with deviant ratings are perceived as manipulated, given that they differ from the majority opinion. Thus, the first research question I pose is: *What is the effect of reviews with deviant ratings on perceived review manipulation?*

Figure 1 Example of negative individual rating

The effect of review content in helping customers evaluate reviews has been demonstrated in previous studies (Aerts et al., 2017; Qi et al., 2016; Willemsen et al., 2011). Customers may be interested in learning details about others' consumption experiences before making their own decisions (Wang et al., 2015); however, not all reviews contain detailed information. Thus, the impacts of detailed (i.e., concrete) information on customer perception is a topic worthy of investigation. Prior studies have suggested that compared to abstract content, concrete content triggers favorable attitudes toward reviewers and perceptions of review helpfulness (Aerts et al., 2017; Schellekens et al., 2010), and the joint effects of concrete review content and other factors have also been identified (Shin et al., 2019; Huang et al., 2018). However, as yet, it is unclear whether content concreteness helps customers evaluate reviews with deviant ratings from the manipulation perspective. Thus, I propose a second research question: *How does content concreteness influence customer perceptions of manipulation toward reviews with deviant ratings?*

Some platforms such as Yelp provide a reviewer profile page that presents rating distributions based on a reviewer's prior ratings in order to assist customers in judging comments and making decisions. The example is presented in Figure 2. If a distribution reflects a reviewer's specific tendency, it can be used to evaluate subsequent ratings (Gao et al., 2017). When others view these profiles, their perceptions of rating distributions can color how they read reviews posted by the specific reviewer. For instance, a reviewer with a positively skewed distribution may be more likely to write negative reviews in general. Thus, if that reviewer posts a negative review, especially if it contradicts majority opinions, other consumers may feel the review is misleading, manipulated, or untrustworthy. Although some aspects of reviewers have been identified as determinants of review perceptions (Filieri et al., 2018; Ngo-Ye & Sinha, 2014), research focusing on reviewer rating distribution is rare (Fang et al., 2016). To fill this gap, the last research question of this study asks: *How does reviewer rating distribution affect customer perceptions of manipulation regarding a review with a deviant rating?*

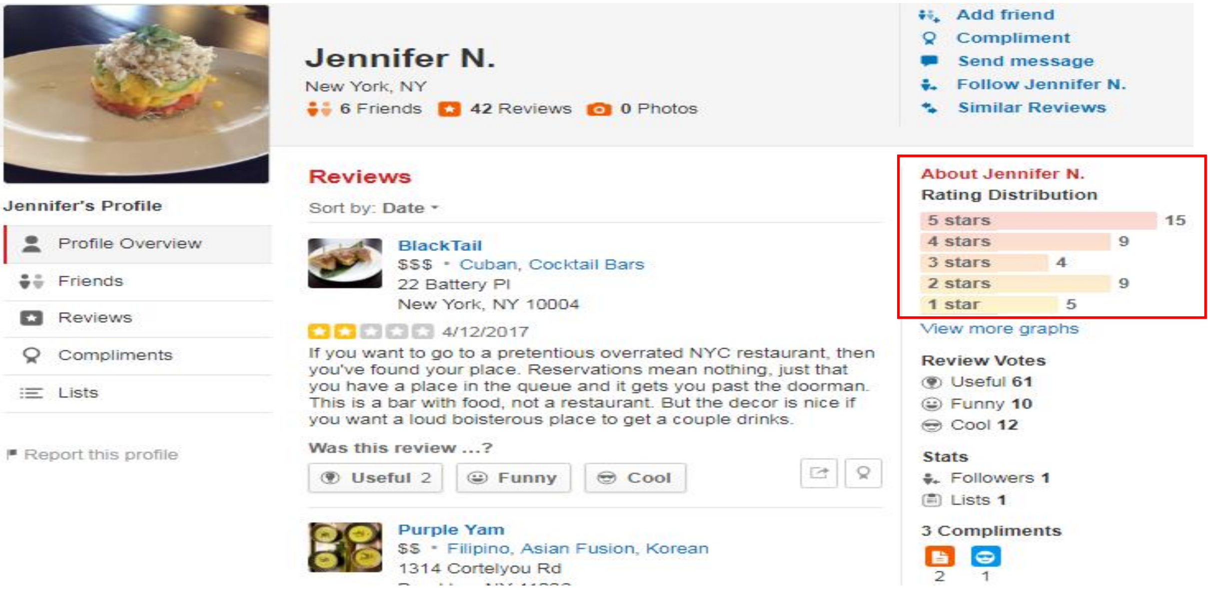


Figure 2 Example of reviewer rating distribution

To explore these questions, this study conducted two online randomized experiments using Amazon Mechanical Turk and one field study using Yelp data. The main experiment was conducted to investigate the effect of rating deviation on perceived review manipulation and identify the moderating role of review content concreteness and reviewer rating distributions. The results suggest that reviews with deviant ratings are more likely to be perceived as manipulated reviews and that review content concreteness and reviewer rating distributions act as moderators between rating deviation and review perceived manipulation. Specifically, when a review contains concrete content, perceived review manipulation based on its deviant rating will be reduced. While negatively skewed distributions may reduce perceived manipulation regarding a review with a deviant rating, a reviewer's positively skewed rating distribution may make ratings with deviant reviews by that reviewer more likely to be perceived as untrustworthy. The complementary experiment explored the influences of rating deviation and perceived review manipulation on indicators that are directly related to customer purchase decisions (perceived review helpfulness/adoption). The results indicate that reviews with deviant ratings are perceived as less helpful and adoptable and suggest that the effects are mediated by perceived review manipulation. The field study using Yelp data validates the research questions tested in the main experiment. The results suggest that reviews with deviant ratings are more likely to be filtered as manipulated reviews. When reviews contain subjective content and the reviewers have larger skewness based on their past ratings, the review filtering is more likely to happen.

In the following sections, I present an overview of prior literature and develop our hypotheses. I then describe the two experiments and one field study with the summary of our results. I conclude with a discussion of the findings and implications.

2.2 Literature Review

2.2.1 Literature on review helpfulness and adoption

Online customer review is conceptualized as "peer-generated evaluations posted on company or third-party websites" (Mudambi & Schuff, 2010). To be specific, it refers to any positive, neutral, or negative opinions about a product or service created and published on any website by a potential, former, or actual customer (Filieri, 2015). It is considered as a new source of information through which customers can interact and exchange shopping experience with each other (Hu et al., 2008). Similar to traditional word of mouth, online review (an electronic word of mouth) is a determinant of customer product choice (Senecal & Nantel, 2004) and purchase intention (Erkan & Evans, 2016). Its importance and quantity urge more research to focus on identifying helpful reviews and insisting customers to make better decisions.

Existing studies have identified several definitions about review helpfulness. One stated that review helpfulness is an index that can reflect how helpful the community found the review and the degree to which other customers believe that the review is helpful (Baek et al., 2012). A similar definition suggests that review helpfulness is the extent to which product evaluations among peers can help and facilitate others' purchase decision process (Mudambi & Schuff, 2010). As an indication of product quality (Schindler & Bickart, 2012), helpfulness vote can attract customers and influence product sales, especially for new published products (Cui et al., 2012).

Review helpfulness can be influenced by several factors. Mudambi and Schuff (2010) found that review extremity and depth of content both affect review helpfulness, and the effect largely depends on the type of products (experience good vs. search good). Ghose and Ipeirotis (2010) showed the impact of review content on review helpfulness. That is, more accurate linguistic and readable text are associated with higher review helpfulness. Sentiment expressed in review text also works that stronger sentiments can enhance review helpfulness (Mousavizadeh et al., 2015). In addition, reviewer-related features can also influence review helpfulness. For instance, Lee and Choeh (2016) found that when a reviewer has higher reputation and disclosures more identity information, his/her review will be perceived more helpful. In general, customers tend to evaluate review helpfulness from two perspectives—the credibility of a review, represented by customers' perception of review source (Chaiken, 1980), and the information expression in a review, that is, review argument quality. Reviewer information including reviewer reputation and expertise are all indicators to infer source credibility (Racherla & Friske, 2012; Lee & Choeh, 2016). Concerning to argument quality, review objectiveness and completeness are both crucial since complete reviews can provide more usable information and objective reviews contain more attribution-related features (Dellarocas, 2003; Wang et al., 2016). In addition, longer and more readable reviews are positively associated with argument quality by containing more persuasive information (Korfiatis et al., 2012; Huang et al., 2015).

Information adoption is the extent to which people accept content after assessing its validity and this definition comes from an individual-level information processing perspective (Goodman & Darr, 1998; Zhang & Watts, 2008). It is an extension of information usefulness that useful information is more likely to be adopted (Sussman & Siegal, 2003). Information adoption in online review suggests the extent to which consumers modify their behavior by utilizing the suggestions made in the reviews (Sussman & Siegal 2003; Cheung et al., 2008). That is, after a customer reads an online review, he/she may choose to accept the opinions from the review to decide whether purchase this particular product or select another product.

Several studies have identified the factors that can influence the information adoption (Zhang & Watts, 2008; Papathanassis & Knolle, 2011; Shen et al., 2016; Hussain et al., 2017). For example, Zhang and Watts (2008) found that argument quality and source credibility are both determinants of information adoption in online reviews based on heuristic-systematic model of information processing. Later study added adoption readiness as another stream of influential factors on information adoption. Adoption readiness is represented by content richness and content accessibility, constituting a precondition of involving in with review information (Papathanassis & Knolle, 2011). In addition, herding factors (i.e., discounting own information and imitating others) play an important role on information adoption (Shen et al., 2016). When a consumer discounts his/her own information, he/she may turn to alternative information. Online review is a good information source. While if a consumer likes to imitate others, he/she can better learn from others' opinions embedded in online reviews. Both increase the probability to adopt the information in online reviews.

2.2.2 Literature on review extremity and deviation

Reviews can be classified to extreme and moderate reviews based on the intensity in review sentiment. One- and five-star ratings indicate extremely negative reviews and extremely positive review, respectively. A review with three-star rating can be considered as a moderate review. For extreme reviews, Mudambi and Schuff (2010) indicated that negative reviews are perceived more helpful relative to positive reviews because customers attribute some non-product reasons to positive reviews. This finding is only applicable to utilitarian products (Sen & Lerman, 2007). If a product is hedonic, its positive reviews are more acceptable.

Concerning to the comparison between extreme and moderate reviews, extreme information posed greater weight on individuals' impression formation, thus reviews with extreme star ratings are more influential on customer perceptions and sale predictions than moderate ratings (Skowronski & Carlston, 1989; Chevalier & Mayzlin, 2006; Forman et al., 2008). While the effect reverses if the product belongs to experience goods (Mudambi & Schuff, 2010). Salehan and Kim (2016) did find that reviews with neutral sentiment are more acceptable since extreme reviews seem less rational when evaluating experience goods.

Social influence suggests a possibility that individuals revise their own behavior to be similar with others (Jahoda, 1959). In online context, customers may also adjust their product evaluations to conform to peers in order to behave accurately or receive social recognition (Deutsch & Gerard, 1955; Price et al., 2006). This reminds scholars to examine rating extremity from the perspective of deviation from others' rating. First stream of research focused on the generation of deviant or consistent reviews. In other words, what factors influence individuals' tendency to write deviant reviews. For example, Hong et al. (2016) suggested that individuals from a collectivist culture such as China and Georgia are less likely to give deviant ratings compared to those from a individualism culture such as U.S and New Zealand since people from collectivist cultures tend to suppress their emotions during the communication with others (Butler et al. 2007).

Second stream of research regarded with the conflicting effects of review deviation on customer perceptions. For example, Danescu-Niculescu-Mizil et al. (2009) showed that deviation from average rating could significantly affect review perception (less review acceptance). There are two possible reasons for this phenomenon. The first may be that the existence of conflicting average rating reduces individual reviews' credibility and diagnosticity (Qiu et al., 2012). That is, the existence of conflicting rating has negative impact on customers' product-related attributions of the review and then decrease review diagnosticity. Another may stem from individuals' general preference for confirmation. This phenomenon is called confirmation bias that individuals tend to perceive reviews that confirm (versus disconfirm) their initial beliefs as more helpful (Yin et al., 2016). Thus, average rating helps customers form initial impression and then a deviant rating suggests an inconsistency from initial impression, lowering perceived acceptance of this comment (Klayman & Ha, 1987; Bao & Chau, 2016). Some research also focused on the positive impact of review deviation. One perspective to explain the positive influence is customer attention. Specifically, a review with larger deviation highlights itself from other reviews and gets itself more attention, increasing its perceived adoption and helpfulness (Shen et al., 2015; Gao et al., 2017). Another possible explanation may come from customers' particular experience from deviant rating. For example, a review deviant from others' ratings may reflect the reviewer's special experience about a product, expressions of his/her own opinion, thus is assessed as more informative by the readers (Forman et al., 2008; Gao et al., 2017).

2.2.3 Literature on review content

As the development of big data techniques, the focus on online reviews has a significant transformation from traditional form of data to novel form of data. Traditional data includes overall review-related information, such as review ratings (also called "review valence") and review number (also referred to "review volume"). Research on the effect of these traditional metrics such as review ratings have mixed findings. For example, several studies found that review volume is more important to product sales compared to review valence (Duan et al., 2008), while others suggest a more prominent role of review valence (Chevalier & Mayzlin, 2006). These inconsistent findings highlight the importance to go beyond traditional data to investigate the novel form of data.

The new form of data is called unstructured data, which contains textual data and non-textual data (e.g., audio and images). Prior studies have started to understand unstructured texts in online review content including customer complaints, experiences, and satisfaction (Lee & Hu, 2005; Xiang et al., 2015). One research stream focused on the generation of different review content. For example, customers at the lower tier properties are more likely to share opinions about transactions and values than those who stay at the middle or higher tier properties (Han et al., 2016). In addition, social network integration also works on linguistic features in review text. Specifically, when a platform integrates with Facebook, its review text has higher positive emotion and lower disagreement expressions (Huang et al., 2017).

Another stream refers to specific dimensions of textual content of online reviews (i.e., statistical and narrative characteristics) (Lee et al., 2008; Qazi et al., 2016; Felbermayr & Nanopoulos, 2016; Hong & Park, 2012; Zhou et al., 2018). Statistical characteristics of review contents contain review length, correctness (Ghose & Iperiotis, 2011), sentiment (Felbermayr & Nanopoulos, 2016), density and diversity (Qazi et al., 2016) and most of these characteristics can influence perceived review helpfulness or product sales. For example, Qazi et al. (2016) identified the number of concepts contained in a review and found that this number can significantly affect perceived helpfulness of the review. For narrative characteristics of review content, although the studies are not as abundant as those in statistical characteristics, they can also play determinant role on customer attitude and purchase decisions (Hong & Park, 2012). Several factors such as culture may contribute to different effects of narrative characteristics of review content. Specifically, Chinese tend to refer to seller trustworthiness, product functionality, price, product quality and product aesthetics, while Americans care more on emotional attitudes and recommendation expressions in online reviews (Zhu et al., 2017).

2.2.4 Literature on reviewer characteristics

Source credibility theory suggests that the perceived credibility of a source can affect communication persuasiveness (Hovland & Weiss, 1951). Source credibility is defined as recipients' perceptions toward a message source (Chaiken, 1980). It is a determinant of information acceptance by positively relating with information acceptance possibility (Zhang & Watts, 2008) and changing customer attitude and information diagnosticity (Ayeh et al., 2013; Filieri et al., 2015).

In online review context, the source of a review suggests the reviewer. Since large amount of online reviews flood into the market, customers need to pick a fraction of reviews to read and at this time, making reviewer characteristics one important factor to decide which reviews to read. Thus, focusing on reviewer characteristics is of great concerns. Prior studies have already begun to pay attention to reviewer characteristics such as reviewer reputation (Otterbacher, 2009), reviewer identity disclosure (Baek et al., 2012), reviewer social ties (Yin et al., 2014), reviewer origin (Lee et al., 2018), reviewer historical rating distribution (Fang et al., 2016), and reviewer influential level (Malik & Hussain, 2018).

Studies about reviewer characteristics mainly fall into two streams. One stream intends to identify the effect (i.e., direct or moderating effect) of reviewer characteristics on perceived review usefulness and product sales (Banerjee et al., 2017). For example, disclosing identity-relevant information about reviewers such as real name (Forman et al., 2008), origin (Filieri et al., 2019) and "real" photo (Park & Nicolau, 2015) can shape other customers' judgement of the reviews. Reviewer-related characteristics also have indirect influence on review perception. Filieri et al. (2019) highlighted a finding that the local origin of a reviewer moderates the relationship between extreme rating and review helpfulness. Specifically, extreme ratings from local reviewers are perceived as more helpful.

Another stream of research shows more interest into the prediction efficiency of review helpfulness by incorporating reviewer characteristics (Ngo-Ye & Sinha, 2014; Zhang et al., 2016; Lee et al., 2018; Malik & Hussain, 2018). For example, Ngo-Ye & Sinha (2014) proposed reviewer's RFM (recency, frequency and monetary value) to characterize the reviewer's overall engagement and improve the prediction accuracy of online review helpfulness. Zhang et al. (2016) further treated reviewer characteristics as nonverbal behaviours and suggested that fake review detection model with these nonverbal behaviours has higher performance. Later study developed a comprehensive research model based on signalling theory and found that the usage of reviewer-related signals (i.e., reviewer expertise) can increase the model performance to predict the most-helpful reviews (Siering et al., 2018).

2.2.5 Theoretical foundation – review manipulation

Online disinformation is defined as "false information that is purposely spread to deceive people" (Lazer et al., 2018). It is regarded as a major problem in the dissemination of information in many important societal realms, including politics, medicine, and commerce. Different from online misinformation, which is a broader concept and describes false claims without considering the falsehood motivation, online disinformation only refers to fabricated information from deliberate intention (Shin et al., 2017). In ecommerce context, the most common pattern of manifestation of online disinformation is manipulated reviews as these reviews are deceptive with the intention of misleading consumers in their purchase decision-making (Hu et al., 2012; Zhang et al., 2016). These deceptive reviews are misleading because the content does not reflect a truthful account of an actual consumer's experience.

Current studies on review manipulation mainly focus on its detection (Akogl et al., 2013; Zhang et al., 2016), influences on firm profits (Hu et al., 2012; Anderson & Simester, 2014), and factors that may impact individuals' incentives to manipulate reviews (Mayzlin et al., 2014; Luca & Zervas, 2016). Since review manipulation has become a relatively common phenomenon, we ask: *Might the examination of review manipulation provide a novel direction for understanding consumers' perceptions of online reviews?* This question lies at the core of our research and is an integration of the three research questions proposed in *Introduction*

To investigate review manipulation from the perspective of consumers' perceptions, our study proposes the definition of perceived review manipulation as the extent to which an individual perceives a review is manipulated. A related concept is perceived review credibility, which refers to the extent to which a consumer perceives the opinions presented in a review as believable, true, or factual (Cheung et al., 2009) and is a key determinant of consumer attitudes toward a service, product, or firm and can influence review adoption (Lim & Van Der Heide, 2014; Hussain et al., 2017). There are three differences between perceived review credibility and perceived review manipulation. The first difference is the valence to consider consumer's trust. That is, perceived review credibility refers to consumer's trust level toward a review from a positive perspective; while perceived review manipulation focuses on consumer's trust level from a *negative* perspective. The second difference regards to consumer motivation. Perceived review credibility only refers to the reliability or the quality of review claims without considering reviewer's motivation, thus a review with low perceived credibility may belong to a type of misinformation. However, perceived review manipulation is mainly about reviewer's specific motivation to mislead potential customers as it is a type of online disinformation. The last difference is the consequence. A review with low credibility may lead customers to make low-quality decisions (i.e., non-optimal product choice). While a review that is manipulated may mislead customers to make contrary decisions (i.e., worst product choice).

2.3 Hypotheses development

Information processing theory suggests that individuals have cognitive structures that store their beliefs, attitude, value, and preference (Rokeach, 1973). Under such cognitive structures, individuals will organize their thinking and experience a range of process after which they undertake a behavior or make a choice. During this process, individuals tend to form expectations, which may direct their attention to relevant information and guide their further evaluation (Entman, 1989; Hann et al., 2007). In this context, individuals may have initial beliefs toward a review with deviant rating and then process other information such as review content and reviewer rating distribution to further evaluate the review.

2.3.1 Review deviation and review manipulation

Review consistency indicates the extent to which the evaluations in a single review are consistent with the evaluations of other contributors toward the same product or service (Zhang & Watt, 2008). Reviews with different consistency (Consistent reviews versus. Deviant reviews) may receive different attention from customers, suggesting that scholars should examine customer perceptions toward individual reviews from the perspective of review consistency/deviation.

Given the rapid growth of online reviews and the lack of uniform reviewing standards, some reviews may be of low quality or may even be posted by fake customers. Since review consistency can influence individual perceptions, the assessment of review authenticity is becoming increasingly important. Individuals tend to believe majority opinions; reviews that differ from the product/service evaluations of most other users are likely to be perceived as less credible and more misleading (Cheung et al., 2009; Zhang & Watt, 2008). Readers may even think that such reviews are written by fake customers; reviews with deviant ratings may be perceived as arbitrary and the opinions expressed may be perceived as more emotional and less objective, leading review readers to judge the review as being of low quality (Hong et al., 2016). These negative perceptions may influence customers’ attitudes toward the review in that customers will be more likely to regard the review as misleading and unreliable (Ayeh, 2015). Based on the definition of perceived review manipulation proposed in the above discussions, the review will be more likely to be perceived as manipulated. Thus, we hypothesize:

H1: Reviews with deviant ratings are more likely to be perceived as manipulated.

2.3.2 Content concreteness and review manipulation

The content of a message can be categorized as concrete or abstract depending on the extent to which a message reduces the guesswork required by the reader (Richardson, 1980). A concrete expression indicates some factual description of an object or event, while an abstract expression is an ambiguous description of an object or action that leaves room for interpretation (Schellekens et al., 2010; Hansen & Wänke, 2010). In the context of online reviews, a concrete review contains detailed information about a customer’s experience and uses specific expressions such as “the cake is too sweet,” whereas an abstract review provides general statements about a product or service (e.g., “the restaurant is good”). Compared to abstract reviews, concrete reviews are perceived as more diagnostic because they contain detailed information and thus reduce uncertainty (Huang et al., 2018; Schindler & Bickart, 2012), and research has suggested that concrete reviews are perceived as more helpful than abstract reviews (Shin et al., 2019).

Concrete opinions in a review with deviant ratings may affect customer perceptions. Because concrete review content provides detailed information about a product or service (Huang et al., 2018), it is often interpreted as objective and diagnostic guidance, and thus increases the perceived quality of the review (Martin et al., 2014). High levels of content concreteness also imply that a reviewer has the ability and desire to express his or her specific opinions about a product or service (Pan & Zhang, 2015). This may prevent customers from judging the reviewer as a fake customer and increase his or her perceived credibility. While deviant ratings are often perceived as misleading, deviant reviews that include concrete content will likely seem less misleading because review readers will be more likely to perceive the review as high quality and the reviewer as credible. Thus, we hypothesize:

H2: Individuals' will be less likely to perceive reviews with deviant ratings as manipulated when the reviews have concrete rather than abstract content.

2.3.3 Reviewer habit and review manipulation

Habit is the behavioral tendency to repeat responses in steady supporting contexts (Ouellette & Wood, 1998). It can be activated in an autonomous process through frequent repetition of an activity in daily life (Evans & Stanovich, 2013; Wood et al., 2002). Personal characteristics such as age, gender, and culture may guide individuals' daily activities and repetition of such activities may cause habits to form. As personal characteristics have been shown to be influential in determining individuals' product evaluations (Jani & Han, 2014; Chung & Darke, 2006), individuals' evaluation tendencies may gradually become habitual. It has been shown that past rating bias has a positive relationship with the rating bias of subsequent reviews (Gao et al., 2017). Thus, individuals' past rating habits may be used as a context for understanding subsequent reviews. Reviewers may write reviews according to habit because making decisions based on habit requires little cognitive effort (Bond et al., 2008). If review readers notice the potential influence of rating habits, their perceptions about the review may change and they may perceive the reviewer as less conscientious and appraise the review as less helpful. In contrast, a review that departs from a reviewer's previous rating tendency may cause others to believe that the reviewer did expend cognitive effort to write the review, resulting in a review that is more likely to be judged as high-quality (Yin et al., 2014).

A few electronic platforms (e.g., Yelp) have designed a special index—a rating distribution, which presents a reviewer's previous reviews according to star-level (i.e., 1-5-star ratings). The distribution of a reviewer's ratings can help customers understand a reviewer's rating habits, which they may take into consideration when judging the review for the purpose of making a purchase decision. For a reviewer with a positively skewed rating distribution, customers might conclude that the reviewer habitually writes negative reviews. If the reviewer contributes a review that is negatively deviant from the average rating (i.e., the majority of reviewers give positive ratings), customers may believe that the reviewer is simply in the habit of giving negative ratings and even conclude the reviewer as a professional poor evaluation blackmailer. Then they may be less likely to believe that the review is credible and tend to suspect that the review is misleading or manipulated. However, if a reviewer has a negatively skewed rating distribution and writes a negatively deviant review, others may perceive the reviewer a departing from his or her usual habit of giving positive reviews and they may judge the review as highly credible and not likely to be misleading or manipulated. Thus, we hypothesize:

H3a: Reviews with negatively deviant ratings are more likely to be perceived as manipulated when rating distributions of the reviewer's past ratings are positively skewed.

H3b: Reviews with negatively deviant ratings are less likely to be perceived as manipulated when rating distributions of the reviewer's past ratings are negatively skewed.

The main research framework is presented in Figure 3.

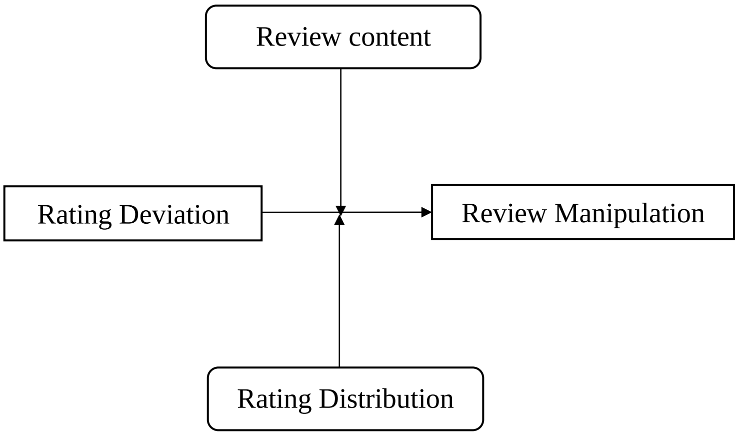


Figure 3 Research framework

2.4 Study 1—Main experiment

To test our hypotheses, we conducted two separate studies, Study 1 (discussed in this section) and Study 2 (discussed in the next section). Both studies were online randomized experiments investigating the perceived manipulation of reviews. Overall, the two studies are similar in experimental design (i.e., variable manipulation/measurement and operation sequence) but have one major difference: the dependent variable (perceived review manipulation in Study 1 versus perceived review helpfulness/adoption in Study 2). These differences enable us to test different hypotheses. We begin with Study 1. This experiment is a 2 (reviews with different ratings: deviant ratings vs. consistent ratings) x 2 (review concreteness: concrete content vs. abstract content) x 4 (review distribution based on past reviewer ratings: negative skewness vs. positive skewness vs. neutral skewness vs. bimodal skewness) between-subject design. Please see Table 1 for the groups. This study was conducted to validate H1, H2, and H3.

Table 1 Experimental conditions

		Reviewer rating distribution and review content							
		Positive skewness		Negative skewness		Bimodal skewness		Neutral skewness	
		Concrete	Abstract	Concrete	Abstract	Concrete	Abstract	Concrete	Abstract
Rating deviation	Yes	Group1	Group2	Group3	Group4	Group5	Group 6	Group7	Group8
	No	Group9	Group10	Group11	Group12	Group13	Group14	Group15	Group16

2.4.1 Manipulation

This study focused on the effect of reviews with deviant and consistent ratings. For manipulated reviews with both deviant and consistent ratings, a rectangular field displaying the average rating was clearly presented to study participants. The average score was fixed as a four-star rating, which has been used in prior research as an average rating for online products with positive valency (Qiu et al., 2012). Since five-star average ratings are uncommon because they indicate that almost all reviewers gave a product/service a five-star rating, we found it more realistic to define a positive valence as a four-star rating. In the consistent-rating condition, individual profile ratings were also set as four-star ratings. In the deviant-rating condition, individual profile ratings were set as two-star star ratings. We felt that a one-star rating would be too extreme and that a two-star rating is sufficient to represent rating deviation.

The operationalization of review concreteness is in line with prior studies (Huang et al., 2018). In each situation (regardless of review valency), the attributes mentioned in review content such as location and service are similar. Table 2 presents the operationalized concrete and abstract review content.

Table 2 Concrete and Abstract Review

Concrete review content for positive rating:
The hotel is near to downtown and we only needed to walk five minutes to arrive at the downtown shopping center. The room was relatively clean and quiet for sleeping. The beds are very comfortable and we slept well. The service is good. We arrived at the hotel at midnight, but the staff still welcomed us with drinks and cookies. The staff also were highly efficient and finished our check-in in one minute! Will come again!
Concrete review content for negative rating:
The hotel is far away and we needed to walk 30 minutes to arrive at the downtown shopping center. The room was relatively clean but noisy for sleeping with lots of street noise. The staff have a lousy attitude and didn't even greet us. We even noticed rude stares from the staff who processed our check-out. They were not very efficient. We arrived at the hotel late and were so tired, but the staff was distracted by other work and it took them an hour to finish our check-in! Will never come again!
Abstract review content for positive rating:
The hotel is very near to downtown and within walking distance of our destination. The room was relatively clean and a suitable place for sleeping. We slept well in the room. The service is good. The staff had a friendly attitude that made all of us feel comfortable and they brightened our mood. Besides their excellent attitude, the staff were also highly efficient and processed our check-in quickly! Will come again!
Abstract review content for negative rating:
The hotel is far from our destination and we had to walk for quite some time to get there. The room was relatively clean but was quite noisy for sleeping. The service is bad. Most of the staff have a bad attitude that made all of us very uncomfortable and ruined our good mood. Beyond their bad attitude, they were inefficient. We had to wait a very long time for our check-in to be completed. Will never come again!

Following prior studies (Fang et al., 2016; Rozenkrants et al., 2017), this study includes four groups of review distributions based on the distribution skewness: *bimodal, positive, negative*, and *neutral* skewness, as these four groups cover most distribution situations. All groups contain the same number of total reviews. In the positively skewed distribution group, one-star ratings are strongly dominant, while in the negatively skewed distribution group, five-star ratings are strongly dominant. The bimodally skewed distribution group contains two clusters of ratings: one cluster of five-star ratings and one cluster of one-star ratings, with fewer two-, three, and four-star ratings. The neutrally skewed distribution group reflects a near-normal rating distribution: the majority of past ratings are three-star ratings, followed by four- and two-star ratings, followed by five- and one-star ratings. We included bimodally and neutrally skewed distribution groups to better identify and differentiate the effect of positively and negatively skewed review distributions. Figure 4 illustrates these four distributions groups.

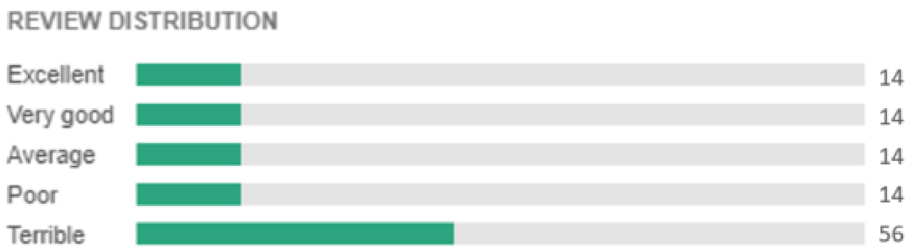


Figure 4a. Positively skewed distribution

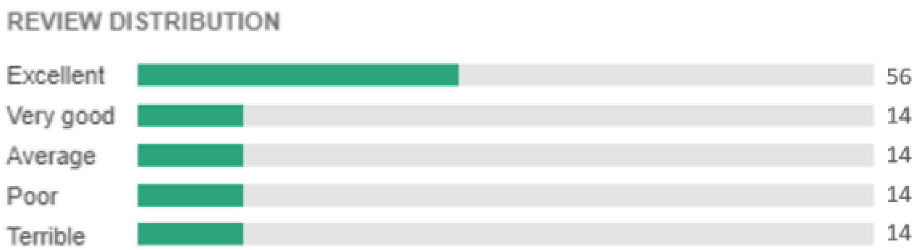


Figure 4b. Negatively skewed distribution

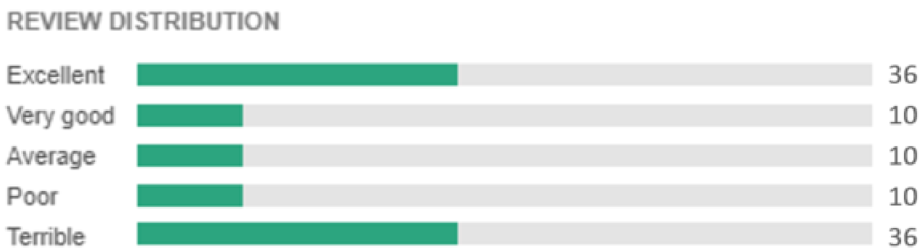


Figure 4c. Bimodally skewed distribution

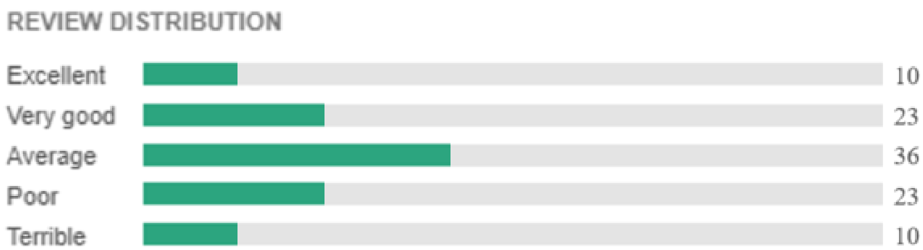


Figure 4d. Neutrally skewed distribution

Figure 4 Four different past review distribution groups

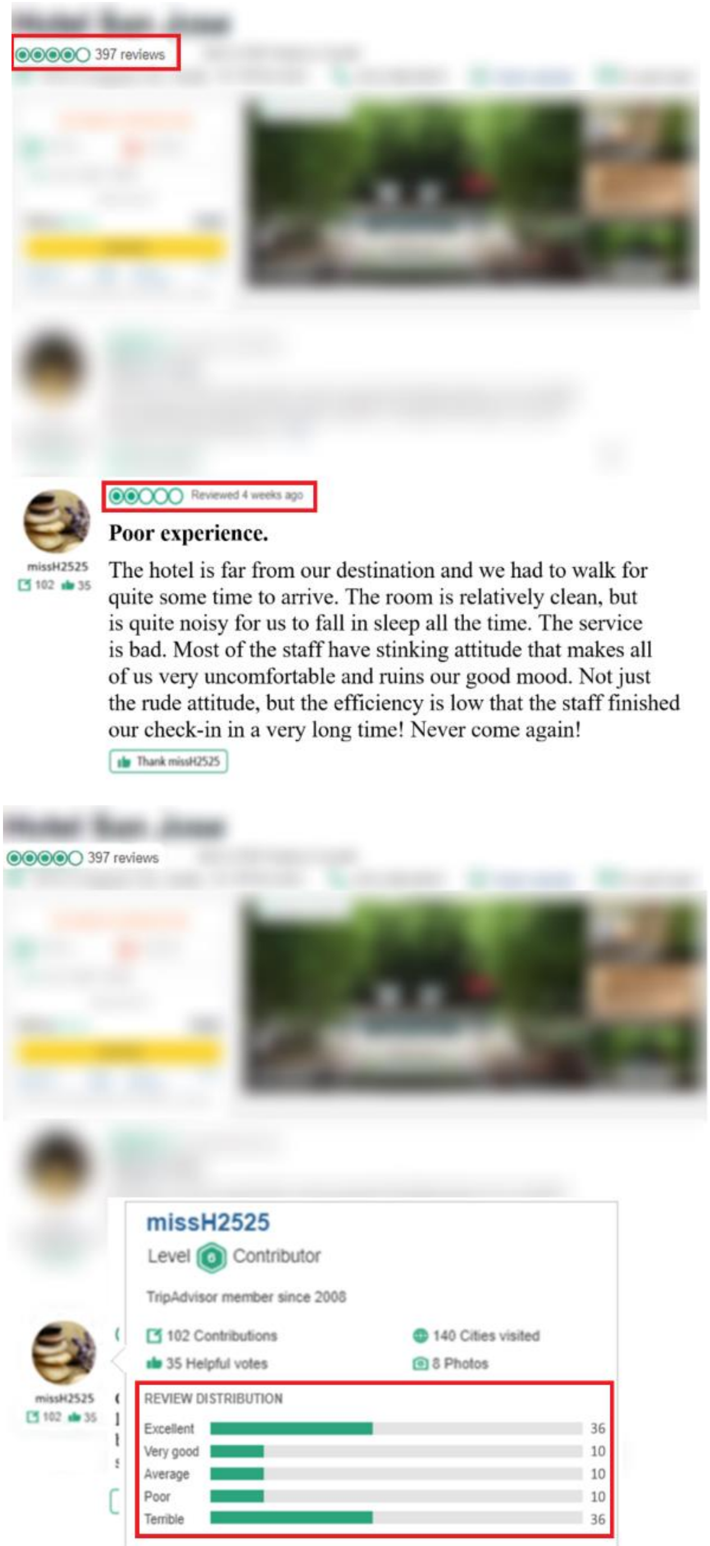
2.4.2 Procedure

The experiment subjects were recruited on Amazon Mechanical Turk (MTurk), which has been established to gather data from diverse populations. Although MTurk has some benefits such as easy access to large sample size, it still has some disadvantages including responder quality and attention distraction in online context (Mason & Suri, 2012). To ensure the experiment quality, we screen the subjects with precautious measures like allowing subjects with HIT approval rates higher than 95% and total number of HITs approved higher than 100 to participate the experiment. We also conduct manipulation checks before formal analysis to check whether subjects pay attention to our manipulations. In addition, our field study with Yelp data showing consistent results may mitigate the disadvantages of running online experiments on MTurk.

After recruiting subjects from MTurk, they were directed to a Qualtrics survey containing images reflecting the different experimental conditions. For each condition, we created two graphic images, which were similar to screenshots from a real online review website; certain elements were blurred using Adobe Photoshop filter-glass tools. Figure 5 shows one condition reflecting a review with a deviant rating, abstract content, and bimodally skewed rating review distribution. After reading all the information on the stimulus image, subjects were asked to respond to questions concerning manipulation checks and dependent variables (perceived review manipulation). At the end, subjects were asked to answer questions related to control variables. After removing incomplete responses, 894 valid subjects remained, yielding 51 to 61 subjects for each condition. Each subject received US$1 for their participation.

2.4.3 Dependent variable

Perceived review manipulation, the dependent variable, is a relatively new concept. We measured this construct by integrating and adapting items from existing studies on manipulated reviews including Mayzlin (2006), Hu et al. (2012), Zhang et al. (2016) (see more details in Appendix). The four items in our scale are (1) “The review is not a truthful account of a real customer’s experience,” (2) “The review intends to mislead customers in their booking decision-making,” (3) “The review is manipulated by related parties (e.g., the hotel itself or the competitor),” (4) “The review is disguised by related parties (e.g., the hotel itself or the competitor).” To test the validity of the item content, we conducted several tests. The Cronbach’s alpha value of these four items is above 0.7. The results of explanatory factor analysis are shown in Table 2 and provide initial evidence for both reliability and validity of this scale. Table 3 provides further evidence of reliability, convergent validity, and discriminant validity with AVEs, correlations, and CRs.



Note: In the left figure, the average rating of the hotel on the upward side is fixed and the rating of the review on the downward side is changed from 2-star to 4-star to manipulate rating deviation. In the right figure, rating distribution is changed to manipulate different conditions of rating distribution.

Figure 5 One condition in the experiment

2.4.4 Control variable

We included several control variables, including review skepticism, which is defined as a basic level of skepticism toward online recommendations. Skepticism results from experiences with prior persuasion attempts and leads to a certain initial level of trust (McKnight et al., 2002); it was measured using five items adopted from Boush et al. (1994): (1) "Online reviews tell the truth," (2) "We can believe what online reviews say"; (3) "The hotels recommended in online reviews are always the best hotels to reserve," (4) "We can depend on getting the truth from most online reviews,"(5) "If an online review were not true, it could not be shown online." All of the above items are reflective indicators and were measured on a seven-point Likert scale from left ("strongly disagree") to right ("strongly agree"). Another control variable measured the usage experience of online review websites by asking subjects: "How often do you search on online review websites (e.g., TripAdvisor or Yelp)" (adapted from Li & Kirkup, 2007). This item was measured on a five-point Likert scale from left ("never") to right ("always"). We adopted the final set of control variables, demographic variables, based on prior studies (e.g., Moore, 2015), and included gender, age, education, and income.

2.4.5 Manipulation check

To ensure the successful manipulation of rating deviation, we asked participants to answer two questions, giving them five answer options: "What is the average rating of the hotel in the above pictures?" and "What is the hotel rating given by the reviewer in the above pictures?" The t-test results showed a nonsignificant difference in average hotel ratings for the deviant-rating group versus the consistent-rating group (for the deviant-rating group, mean = 3.877, S.D. = 0.503; for the consistent-rating group, mean = 3.891, S.D. = 0.443; $t = 0.4369$, $p > 0.1$) and a significant difference of hotel rating by the reviewer between the two groups (for deviant-rating group, mean = 2.230, S.D. = 0.676; for consistent-rating group, mean = 3.837, S.D. = 0.668; $t = 35.750$, $p < 0.05$). The results suggest that the manipulations of rating deviation were successful.

The manipulation check of review content concreteness was measured using a seven-point Likert scale question: "The review makes it easy for you to form mental images of the reviewer's experience" from left ("strongly disagree") to right ("strongly agree") following prior research (Sadoski et al., 1993). We conducted a t-test to compare the measurement for the concrete-content group (mean = 5.637, S.D. = 0.822) versus the abstract-content group (mean = 5.450, S.D. = 0.823). A significant difference was found between these two groups ($t = -3.406$, $p < 0.01$), indicating that the manipulation of review content concreteness was effective.

We performed a manipulation check for different rating distributions by asking subjects to choose an answer from one of the four distributions presented in Figure 2 in response to the question: “What does the review distribution look like?” In the positively skewed rating distribution group, 172 out of 222 subjects chose Figure 4a; in the negatively skewed rating distribution group, 183 out of 225 subjects chose Figure 4b; in the bimodally skewed distribution group, 185 out of 223 subjects chose Figure 4c; and in the neutrally skewed distribution group, 186 out of 226 subjects chose Figure 4d. All ratios were higher than 75%, validating the manipulation of review distributions.

2.4.6 Hypothesis testing

We first ran an exploratory factor analysis and computed Cronbach’s alpha values using two measurement variables (perceived review manipulation and review skepticism). The factor analysis results revealed that both convergent and discriminant validity were good. The value for the last measurement of review skepticism was low, so we deleted this measurement and conducted further checks. All Cronbach’s alpha values were above the threshold of 0.7. The CR and AVE values further provide evidence of good convergent validity. In addition, the square root of the AVE value for each construct exceeds the correlation value between that construct and other constructs, indicating satisfactory validity (Fornell & Larcker, 1981).

H1 posits that reviews with deviant ratings are more likely to be perceived as manipulated reviews. T-test result shows that the mean value of perceived review manipulation is larger in the deviant-rating condition (mean = 3.167, S.D. = 0.062) than in the consistent-rating condition (mean = 2.710, S.D. = 0.062). Thus, H1 is supported; the results are illustrated in Figure 6.

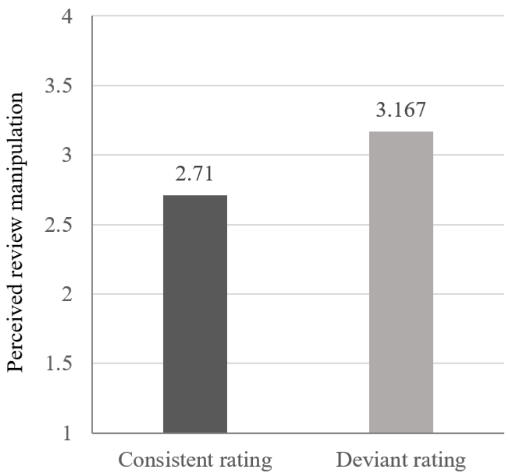


Figure 6 The moderating effect of review content concreteness

We also tested H2, which posits that individuals' perceptions that reviews with deviant ratings are manipulated will be reduced if reviews have concrete versus abstract content. We first employed ANOVA analysis to test whether the interaction between review content and rating deviation exerts a significant effect, which the results confirmed ($F = 10.09$, $p < 0.01$). Then, pairwise comparison was used to investigate the different effects of rating deviation with abstract versus concrete content. The results indicated that subjects in the abstract-content condition had higher manipulation perceptions of deviant-rating reviews (mean = 3.420) than consistent-rating reviews (mean = 2.696); the perception difference was significant ($t = 5.97$, $p < 0.01$). Subjects' perceptions of review manipulation were similar with both deviant (mean = 2.897) and consistent ratings (mean = 2.724) in the concrete-content condition ($t = 1.39$, $p > 0.1$). Thus, H2 is supported; the results are illustrated in Figure 7 and Figure 8.

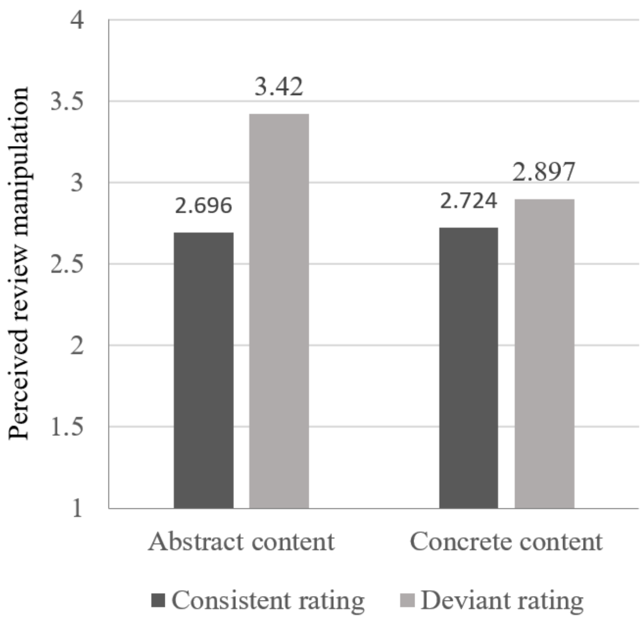


Figure 7 The moderating effect of review content concreteness

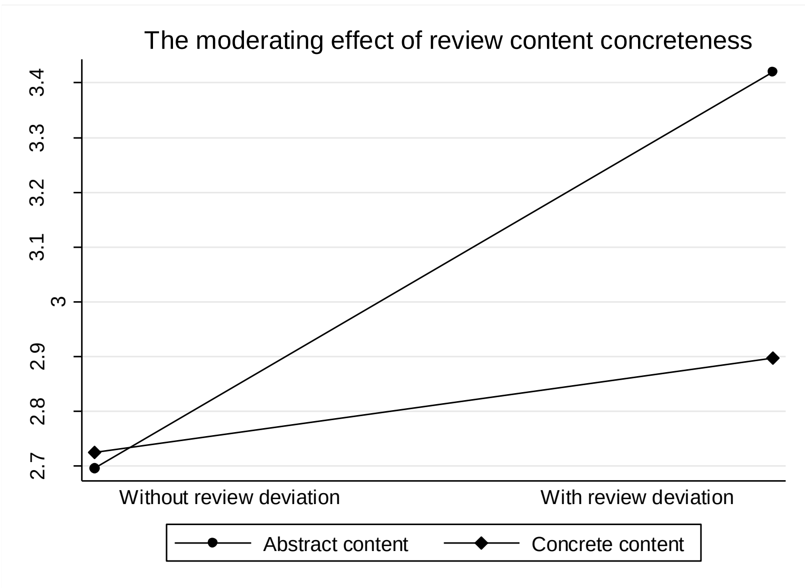


Figure 8 The moderating effect of review content concreteness

This study focused on the moderating role of the reviewer's past-rating distribution to test H3, which predicts that individuals will be more (less) likely to perceive reviews with deviant ratings as manipulated when the reviewer's past rating distribution is positively (negatively) skewed. ANOVA analysis results suggest a significant effect ($F = 5.35$, $p < 0.05$). We thus employed pairwise comparison to further test the effect of different past-rating distributions. The results indicate that subjects in the positively skewed rating distribution condition perceived reviews with deviant ratings as manipulated (mean = 3.692) to a greater extent than reviews with consistent ratings (mean = 2.861); the difference in perceptions was significant (t = 4.78, $p < 0.01$). In the negatively skewed rating distribution condition, perceptions of review manipulation were similar (t = 1.52, $p > 0.1$) across both deviant-rating (mean = 2.894) and consistent-rating groups (mean = 2.630). In the bimodally/neutrally skewed distribution group, the perception of review manipulation did reveal a significant difference between the deviant-rating (mean = 3.071; 3.014) and consistent-rating groups (mean = 2.713; 2.640), but t-test results reveal that the significance of the difference is not as strong as it is in the positively skewed rating distribution condition (t = 2.07; 2.17, $p < 0.1$). We present the results in Figure 9-11, which shows that for reviewers with a positively skewed rating distribution, the effect of rating deviation on perceived review manipulation is stronger than it is for reviewers with other distributions. In addition, the results in Figure 12-13 suggest that, for reviewers with a negatively skewed rating distribution, the effect of rating deviation on perceived review manipulation is weaker than it is for reviewers with other distributions. Thus, H3a and H3b are supported. The general results are illustrated in Figure 14.

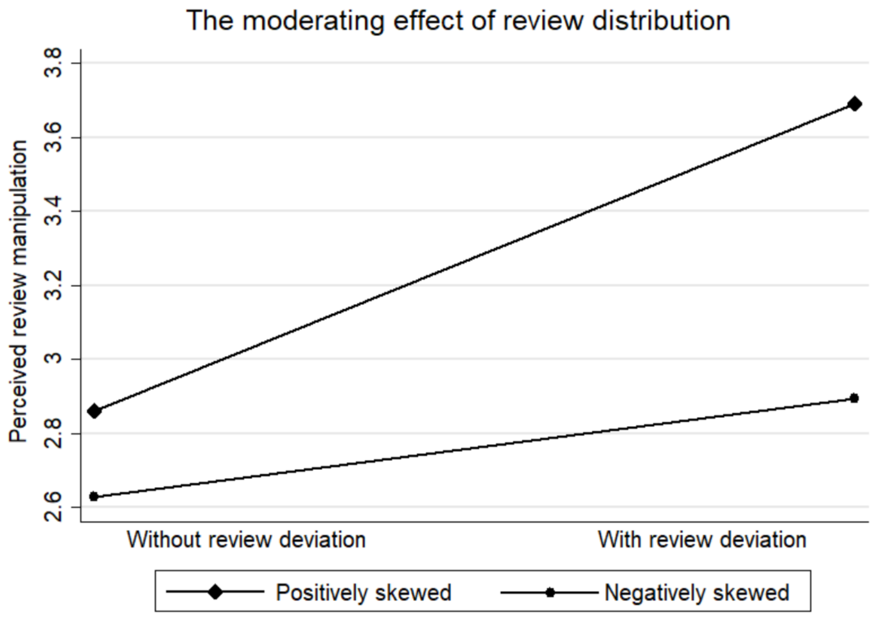


Figure 9 The moderating effect concerning to negatively and positively skewed distribution

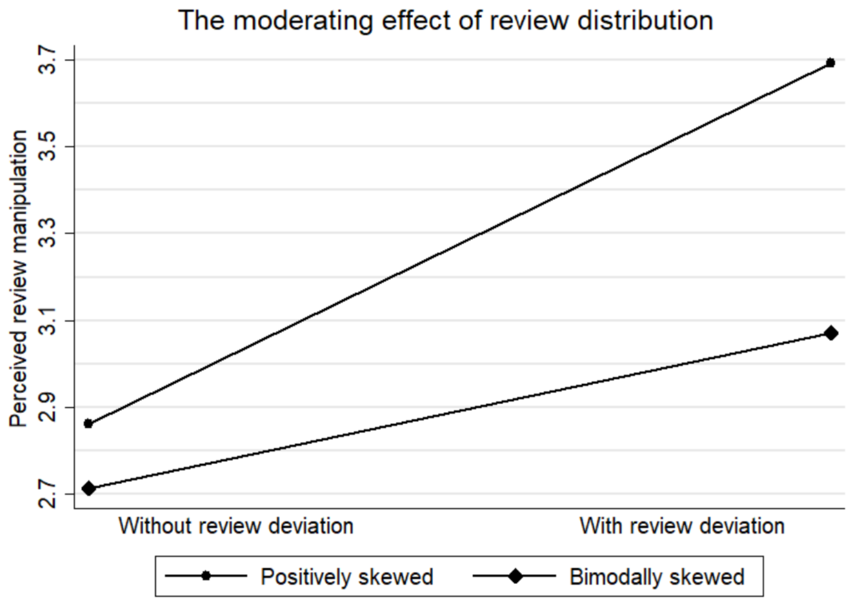


Figure 10 The moderating effect concerning to positively and bimodally skewed distribution

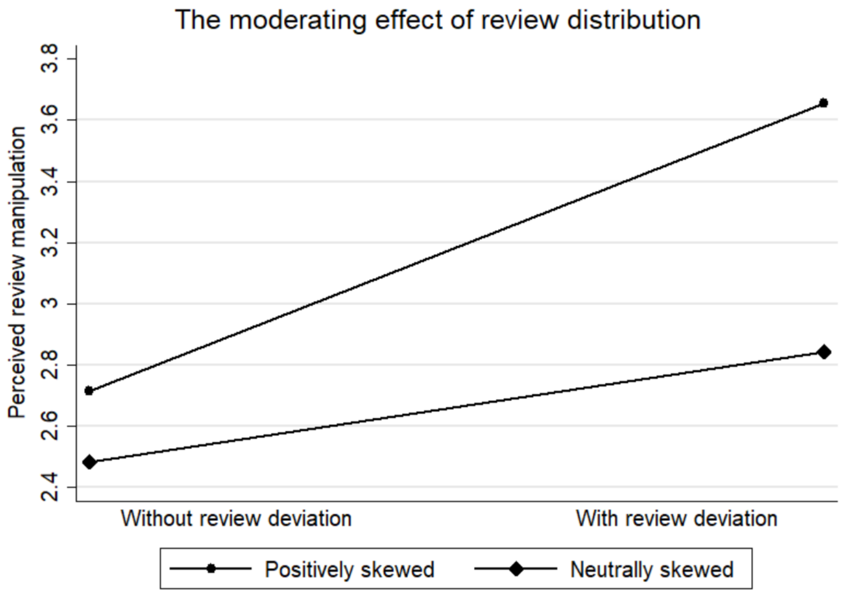


Figure 11 The moderating effect concerning to positively and neutrally skewed distribution

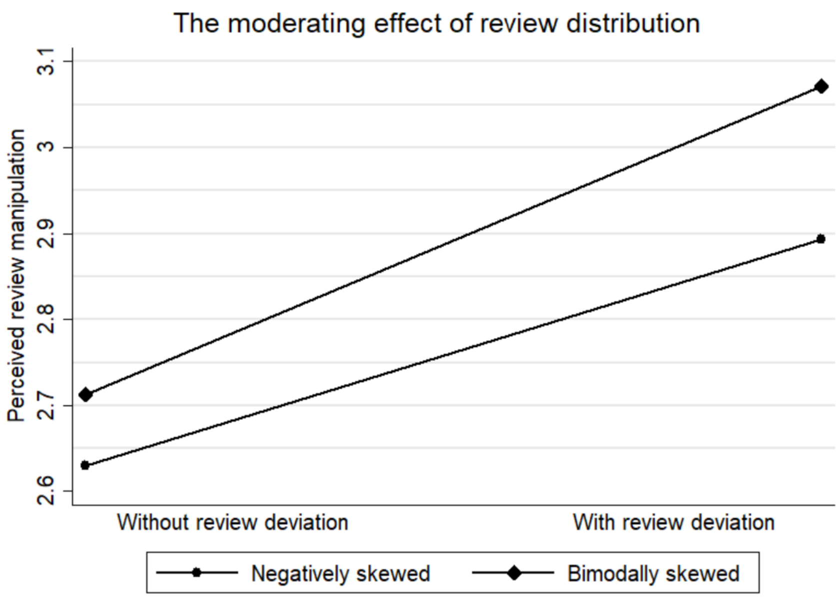


Figure 12 The moderating effect concerning to negatively and bimodally skewed distribution

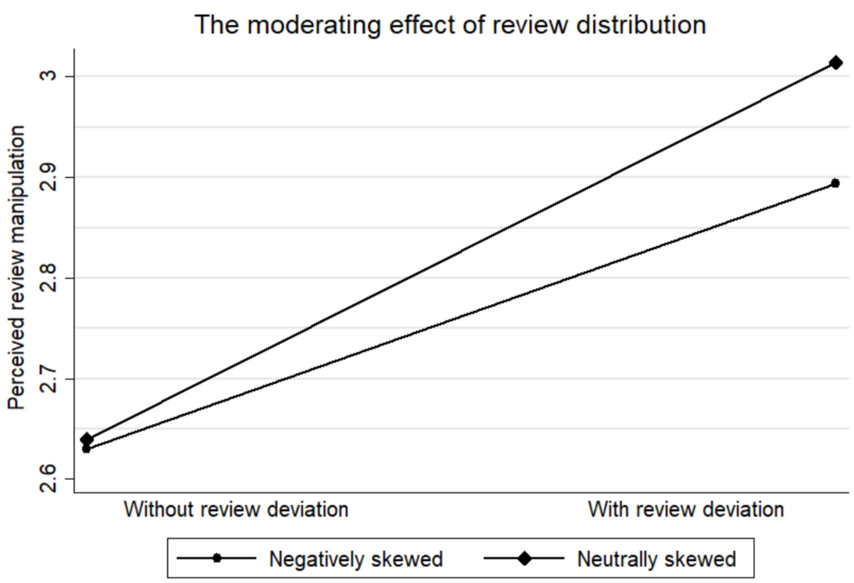
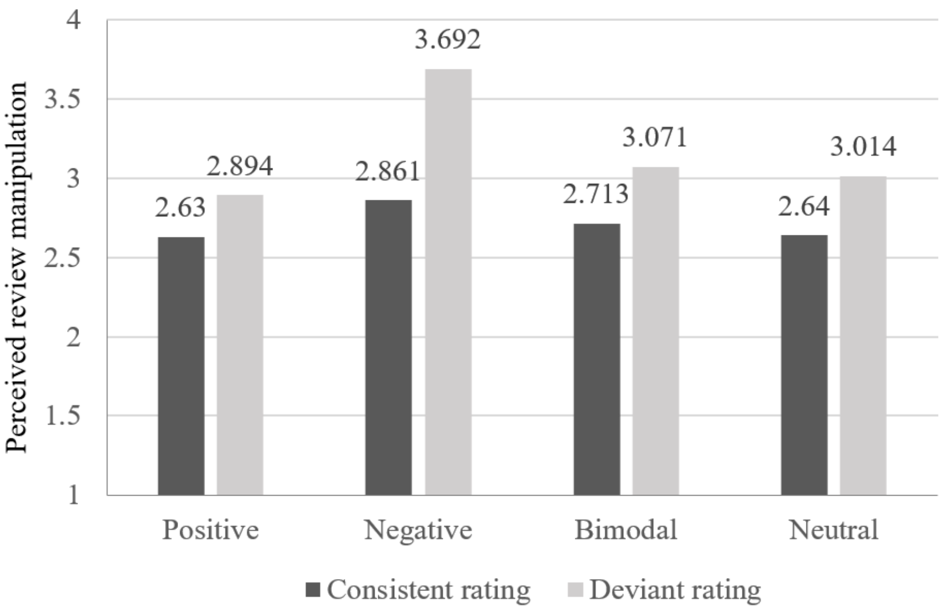


Figure 13 The moderating effect concerning to negatively and neutrally skewed distribution



Note: C represents consistent rating and D represents deviant rating

Figure 14 The moderating effect of reviewer rating distribution

2.5 Study 2—Complement experiment

2.5.1 Effects on decision-making perception

Up to this point, we have discussed how reviews with deviant ratings can influence readers' perceptions of review manipulation, which, in turn, can affect perceptions of review helpfulness and adoption. Given that helpful reviews and review adoption can directly influence purchase decisions (Chen et al., 2014), a better understanding of these two indicators can offer clear benefits to online platforms, firms, and reviewers (Yin et al., 2014). We argue that perceived review helpfulness and adoption are influenced by rating deviation/consistency and perceptions of review manipulation. Review content features (i.e., content concreteness vs. abstractness) and reviewer features (i.e., rating distribution of previous reviews) may have additional effects beyond the rating deviation/consistency and perceived manipulation aspects that we have already hypothesized above.

Confirmation bias denotes the tendency of individuals to overweight the significance of information that confirms (vs. disconfirms) their initial beliefs and positions (Nickerson, 1998). In other words, individuals generally find information that confirms their initial beliefs and thoughts to be more appealing than information that contradicts their initial conjectures (Klayman & Ha, 1987). Thus, customers tend to experience discomfort with reviews with deviant ratings since they differ from the average rating and thus also their initial perceptions based on the average rating (Yin et al., 2016) and may thus form negative opinions about the review, which thereby reduces the adoption likelihood of the review.

In addition to triggering customers' confirmation bias, consistent with attribution theory (Qiu et al., 2012; Kelley, 1967), reviews with deviant ratings also influence product-related attributions of the review. Specifically, when a reviewer evaluates a product differently from others (i.e., gives a deviant review), other customers tend to believe that the reviewer's evaluations are based on non-product-related factors such as personal preference rather than product-related factors, in which case the review may be perceived as less helpful. In contrast, if a consensus exists between the majority's opinions and the reviewer's evaluations, others will attribute the evaluations to product-related factors and treat the reviewer's opinions as usable and helpful, making them more likely to adopt the review in their decision-making (Filieri et al., 2018; Hussain et al., 2017). In other words, a review with a deviant rating that challenges a customer's confirmation bias and appears to be based on attributions unrelated to the product will be less persuasive. Therefore, we hypothesize:

H4a: Reviews with deviant ratings are perceived as less helpful than reviews that align with the average rating.

H4b: Reviews with deviant ratings are less likely to be adopted than reviews that align with the average rating.

Since a review that is perceived as manipulated is judged as containing misleading information and being less credible, the perceived helpfulness and adoption of this review will be reduced (Hussain et al., 2017). The effect of perceived review manipulation can be discussed according to two aspects—reviewer credibility and review argument quality, both of which are determinants of review helpfulness (Aghakhani et al., 2017; Huang et al., 2015; Keller & Staelin, 1987). If a review is perceived as manipulated, others may doubt that the reviewer is an actual customer, thus reducing source credibility. However, review accuracy, an aspect of argument quality referring to the perception that a review properly represents the product information (Delone & McLean, 2003), can also determine the effect of perceived review manipulation. When a review is perceived as manipulated, customers may doubt that the review content conveys accurate product information, thus decreasing perceived review accuracy. Taken together, perceptions of review manipulation reduce perceived reviewer credibility and perceived review argument quality, thus posing effects on perceptions of review helpfulness and adoption. Thus, we hypothesize:

H4c: Reviews that are perceived as manipulated are perceived to be less helpful.

H4d: Reviews that are perceived as manipulated are less likely to be adopted in making purchase decisions.

Next, we conducted Study 2. This experiment was a between-subject design using two different conditions—reviews with consistent ratings versus reviews with deviant ratings—and sought to test the user decision-making indicators posited in H4. The measurements of perceived review manipulation and review skepticism, as well as the other control variables, are identical to Study 1.

2.5.2 Manipulation

This study focuses on the effect of online reviews with deviant and consistent ratings and the effect of perceived review manipulation on perceived review helpfulness/adoption. The manipulation of reviews with deviant and consistent ratings was similar to that of Study 1. In the consistent-rating condition, individual ratings and average ratings were both set as four-star ratings; in the deviant-rating condition, individual ratings were set at two stars.

2.5.3 Procedure

This experiment was also conducted on Amazon Mechanical Turk (MTurk). Like in Study 1, subjects were recruited from MTurk and directed to a Qualtrics survey. For each condition, subjects were presented with one image, a screenshot from a real online review website with identifying elements blurred (e.g., Figure 5, first image). After reading the stimulus image information, subjects were asked to respond to questions concerning manipulation checks and main variables (perceived review manipulation, helpfulness, and adoption). At the end, subjects were asked to answer several questions related to control variables. After removing incomplete responses, 209 valid subjects were left. Among these, 111 were assigned to the deviant-rating condition and 98 were assigned to the consistent-rating condition. Each subject received US$1 for participating.

2.5.4 Dependent variable

The perceived review helpfulness, one dependent variable, was measured using three measurement items adopted from Filieri (2015): (1) “The review is helpful for me to evaluate the hotel,” (2) The review is helpful in familiarizing me with the hotel, (3) “The review is helpful for me to understand the performance of the hotel." Perceived review adoption, another dependent variable, was measured using four measurement items adopted from Filieri (2015): (1) "The review makes it easier for me to make booking decisions (e.g., to book or not to book)," (2) "The review enhances my effectiveness in making booking decisions," (3) "The review motivates me to make a booking decision (or not)," (4) "I will adopt the review's recommendation to book the hotel (or not)."

2.5.5 Manipulation check

This study also asked participants two questions (with five possible answers) regarding the average hotel rating and the average reviewer hotel rating to conduct manipulation checks of deviation and consistent rating. A t-test was used to compare the mean of average ratings of the hotel in the deviant-rating group (mean=3.865, S.D. = 0.041) and consistent-rating group (mean=3.939, S.D. = 0.024); the results revealed a nonsignificant difference ($t = 1.487$, $p > 0.1$). This study also conducted a t-test to compare the mean of the reviewer hotel rating for the deviant-rating group (mean=2.063, S.D. = 0.035) and consistent-rating group (mean=3.908, S.D. = 0.033); the results revealed a significant difference ($t = 38.510$, $p < 0.01$). Thus, a successful manipulation of two groups can be assumed.

2.5.6 Hypothesis testing

We also ran an exploratory factor analysis and computed Cronbach's alpha values on all measurement variables (perceived review helpfulness, adoption, and manipulation as well as review skepticism). The factor analysis results reveal good convergent validity. However, the discriminant validity between perceived review helpfulness and adoption was not sufficient as the measurements of these two variables are similar (shown in Table 3). Since we used these two variables as dependent variables separately, this insufficiency did not influence our results. Since the value for the last measurement of review skepticism is low, we deleted this measurement and conduced further checks. As in Study 1, all Cronbach's alpha values were above the threshold of 0.7, and the results of CR and AVE values provide further evidence of good convergent validity. In addition, the square root of the AVE value for each construct exceeds the correlation between that construct and other constructs, indicating satisfactory validity. The results are presented in Table 4.

Table 3 Factor Analysis of Measurement Variables

	Component		
	Decision-making indicator	**Review skepticism**	**Review manipulation**
RevHelp1	**.901**	.145	-.125
RevHelp2	**.836**	.227	-.098
RevHelp3	**.895**	.145	-.088
RevAdopt1	**.881**	.087	-.153
RevAdopt2	**.878**	.085	-.212
RevAdopt3	**.874**	.107	-.091
RevAdopt4	**.855**	.201	-.032
RevManipu1	-.334	-.031	**.660**
RevManipu2	-.288	-.131	**.720**
RevManipu3	-.006	-.037	**.894**
RevManipu4	-.045	-.068	**.884**
RevSkep1	.162	**.852**	-.113
RevSkep2	.088	**.818**	-.197
RevSkep3	.157	**.748**	.083
RevSkep4	.113	**.823**	-.242
RevSkep5	.162	**.443**	.288

Note: Review Helpfulness= (RevHelp1+ RevHelp 2+ RevHelp 3)/3; Review Adoption= (RevAdopt1+ RevAdopt2+ RevAdopt3+ RevAdopt4)/4; Perceived Review Manipulation= (RevManipu1+ RevManipu2+ RevManipu3+ RevManipu4)/4; Review Skeptical Feelings= (RevSkep1+ RevSkep2+ RevSkep3+ RevSkep4+ RevSkep5)/5.

Table 4 Internal consistency and discriminant validity of constructs

	AVE	CR	*RevAdopt*	*RevHelp*	*RevMani*	*RevSkep*
RevAdopt	0.790	0.938	0.889			
RevHelp	0.771	0.910	0.864	0.878		
RevMani	0.700	0.903	-0.333	-0.316	0.837	
RevSkep	0.684	0.896	0.301	0.338	-0.2224	0.827

H4a assumes that with the existence of deviant ratings, reviews are perceived as less helpful. The t-test result showed a significant influence of rating deviation on review helpfulness ($F = 46.14$, $p < 0.01$). Specifically, the mean value of perceived review helpfulness was smaller in the deviant-rating condition (mean = 4.80, S.D. = 0.117) than in the consistent-rating condition (mean = 5.639, S.D. = 0.124), thus supporting H4a. Figure 15 presents these results.

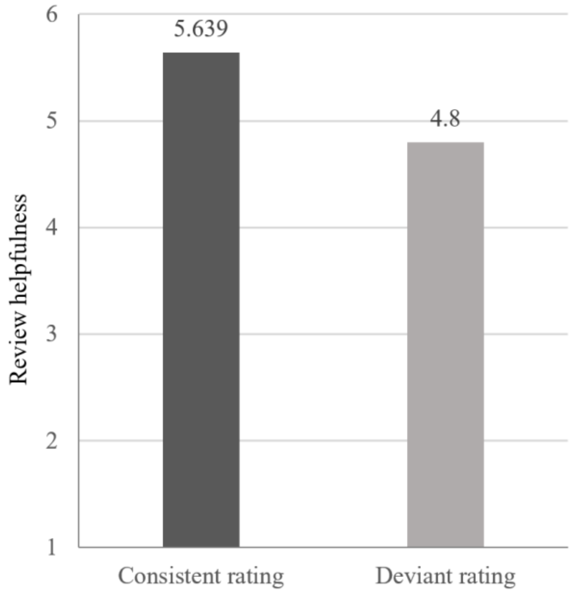


Figure 15 The effect of rating deviation on review helpfulness

H4b posits that reviews with deviant ratings are more likely to be adopted than those with consistent ratings. The t-test result showed that the mean value of perceived review adoption is smaller in the deviant-rating condition (mean = 4.448, S.D. = 0.116) than in the consistent-rating condition (mean = 5.459, S.D. = 0.124), indicating that H4b is supported. The results are shown in Figure 16.

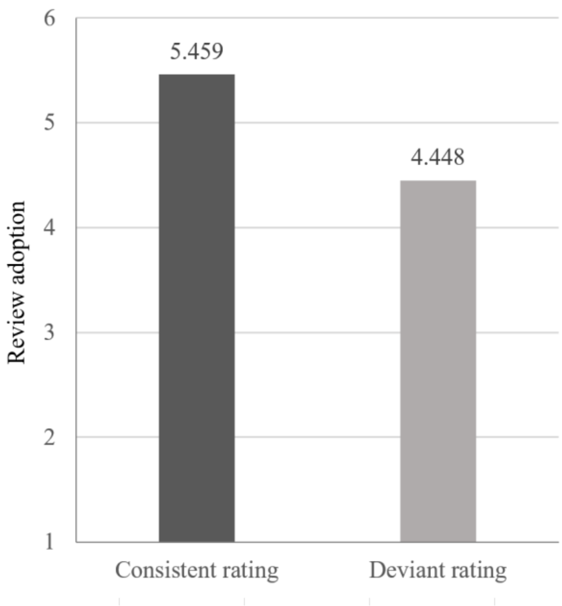


Figure 16 The effect of rating deviation on review adoption

Table 5 Mediation effect using a three-step approach

	Model 1	Model 2	Model 3	Model 4	Model 5
	RevHelp	*RevAdopt*	*RevMan*	*RevHelp*	*RevAdopt*
RatDev	-1.117***	-0.993***	0.494***	-1.000***	-0.857***
	(0.167)	(0.167)	(0.155)	(0.167)	(0.166)
RevMan				-0.236***	-0.276***
				(0.074)	(0.074)
Gender	-0.127	-0.156	-0.020	-0.132	-0.161
	(0.166)	(0.166)	(0.153)	(0.162)	(0.161)
Age	-0.013*	-0.015**	-0.003	-0.014*	-0.016**
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)
Education	0.159	0.103	0.127	0.189	0.139
	(0.122)	(0.122)	(0.113)	(0.120)	(0.119)
Income	-0.022	-0.014	-0.052*	-0.034	-0.029
	(0.030)	(0.030)	(0.028)	(0.030)	(0.029)
RevFeel	0.397***	0.324***	-0.255***	0.337***	0.253***
	(0.091)	(0.092)	(0.085)	(0.091)	(0.091)
WebUse	0.107	0.171*	0.125	0.136	0.205**
	(0.095)	(0.096)	(0.088)	(0.094)	(0.093)
Constant	3.603***	3.764***	3.457***	4.420***	4.719***
	(0.706)	(0.709)	(0.654)	(0.737)	(0.733)
N	209	209	209	209	209
Adj-R^2	0.262	0.217	0.088	0.294	0.265
F	11.53***	9.25***	3.86***	11.81***	10.36***

Note: RatDev—Negative Rating Deviation; RevMan—Perceived review manipulation; RevFeel—Skeptical feelings toward a review; WebUse—Frequency to search on online review website.

H4c and H4d posit the effect of perceived review manipulation on perceived review helpfulness/adoption. To test this hypothesis, we identified the mediating role of perceived review manipulation on rating deviation and perceived review helpfulness/adoption following the three steps by Baron and Kenny (1986), which are widely used in prior studies (Huang et al., 2019; Yin et al., 2014). A mediating effect exists if (1) the effect of the independent variable on the dependent variable is significant, (2) the effect of the independent variable on the mediating variable is significant, and (3) the effect of the independent variable on the dependent variable becomes insignificant or decreases substantially when the mediating variable is included in the model. Table 5 indicates that the effect of rating deviation in Models 4 and 5 is significantly smaller than it is in Models 1 and 2, suggesting the existence of partial mediation. Our study also employed bootstrapping, which is a preferred approach for testing mediation effects because of the loose requirements concerning sample size and distribution (Hayes, 2009; Preacher & Hayes, 2008). The results, based on 1000 replications, are presented in Tables 6 and 7 and further confirm the mediating role of perceived review manipulation in the relationship between rating deviation and perceived review helpfulness/adoption. The existence of this mediating role demonstrates the significant effect of perceived review manipulation on perceived review helpfulness/adoption. Thus, H4c and H4d are supported.

Table 6 Mediation effect on review helpfulness using bootstrapping

	Coef	S.D.	Z	Bootstrapping 95% CI	
				Lower	Upper
RevMan	-0.164	0.062	-2.63	-0.285	-0.042
RatDev	-0.847	0.163	-5.20	-1.167	-0.528

Table 7 Mediation effect on review adoption using bootstrapping

	Coef	S.D.	Z	Bootstrapping 95% CI	
				Lower	Upper
RevMan	-0.149	0.579	-2.57	-0.262	-0.353
RatDev	-1.010	0.158	-6.38	-1.320	-0.700

2.6 Study 3—Field study

The primary goal of Study 3 was to validate H1-H3 by exploring the effects of rating deviation, review content concreteness and reviewer rating distribution on review manipulation. The difference between Study 3 and the first two studies is the dependent variable. The dependent variable in the first two studies is perceived review manipulation, which is a concept proposed from the perspective of customer perception. However, the dependent in Study 3 is review manipulation, which is a relatively objective indicator calculated by Yelp filter algorithm in a real word setting. To validate the hypotheses, we employ data containing Yelp's filtered (manipulated) and unfiltered (non-manipulated) reviews of Chicago restaurants. Yelp was chosen because it uses a filtering algorithm to filter suspicious reviews and puts them in a filtered list. Because the number of helpful votes for a review is highly correlated with the duration that the review is listed in normal page, it cannot reasonably measure customers' perceived helpfulness. Thus, Hypotheses 4 are not tested in this study.

2.6.1 Data collection

We use the original data of Mukherjee et al. (2013), which combines Yelp filtered and unfiltered reviews. To ensure the review credibility, Yelp developed a filtering algorithm in 2005 to identify suspicious reviews. As this algorithm has evolved over these years, Yelp is confident about its accuracy and makes its filtered reviews public. Yelp's filter algorithm has been claimed with high accuracy by a BusinessWeek study (Weise, 2011).

The original data contains 58,517 reviews across 130 restaurants in Chicago. Individual reviews are used as the unit of analysis and every review is marked as Y, N, YR or NR. Reviews that are marked by Y and N are obtained from the restaurant page in which reviews with Y come from the filtered section and those with N come from the regular page. Reviews that are marked by YR and NR are obtained from the reviewer profile page. For any review, Mukherjee et al. (2013) go through all reviews of that particular business. If it is available on regular page, it is given a NR value. Rather, it is given a YR value. Following Mukherjee et al. (2013), we only use reviews with label Y and N. As we are only interested in the reviews with negatively deviant ratings, we keep reviews whose rating is lower than average rating of the involved business. After the data processing, analysis was conducted on a dataset with 8,487 reviews across 103 restaurants.

We calculate the price distribution in our data to see the distribution of restaurant level. Price is transformed from the symbol "$, $$, $$$, $$$$" on the restaurant page with 1 represents $, 2 represents $$, 3 represents $$$ and 4 represents $$$$. Not all restaurants have this price symbol, thus the number of observations of price is lower than restaurant number. The distribution of restaurant price is presented in Table 8, indicating that most restaurants in our data is economic type.

Table 8 The frequency analysis on restaurant price range

	Number	Percentage (%)
$	1,872	26.50
$$	3,840	54.35
$$$	1,101	15.58
$$$$	252	3.57

2.6.2 Variables

The dependent variable of interest, review manipulation, is calculated by Yelp. In each restaurant, Yelp presents a review list called "xx filtered reviews for xx" under which all filtered reviews for this particular restaurant are shown. A review that is included as one of the filtered reviews will be marked as manipulated review; a review that is shown in regular page will be marked as normal review. Therefore, review manipulation is a dummy variable with 1 indicating the review is manipulated and 0 indicating the review is un-manipulated. Table 9 presents the summary statistics of all variables. The average value of the review manipulation was 0.101, indicating that the ratio of manipulated reviews is lower than that of un-manipulated reviews. As the frequency analysis shows (Table 10), among 8,487 reviews, 7,626 are filtered (manipulated) ones and 861 are un-filtered (un-manipulated) ones.

Table 9 Descriptive analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
Manipulation	8487	0.1014	0.3019	0	1
Deviation	8487	0.6433	0.4790	0	1
Subjectivity	8487	0.4935	0.5000	0	1
Skewness	8487	0.2952	1.1460	-3.1870	6.2472
RevLength	8487	137.677	116.769	1	944
WordLength	8487	14.5527	8.1170	1	255
JoinMonth	8487	48.864	20.1064	0	96
Delivery	8487	0.1212	0.3264	0	1
Wheelchair	8487	0.6713	0.4698	0	1
PriceRange	7056	1.9622	.7498	`	4

Table 10 Frequency analysis on filtered reviews

	Number	Percentage (%)
Un-filtered	861	10.14
Filtered	7,626	89.86

The independent variable of interest, review deviation, is a dummy variable. As stated above, we delete reviews whose rating is larger than average rating. We then assign different deviation values for reviews with different rating or same rating. For reviews with deviant ratings from restaurant average rating, deviation is assigned as 1; for reviews with consistent ratings from restaurant average rating, deviation is assigned as 0. From Table 2.9, we can find that the average value of deviation is around 0.6, indicating that the sample is relatively balanced between deviant rating and consistent rating with a little more reviews having deviant rating.

The moderating variable of interest, subjectivity, is measured by calculating subjectivity score in the review text using *TextBlob* package from Python. *TextBlob* is a Python library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more. It has been used in several areas such as marketing and IS literatures (Yoon et al., 2019; Micu et al., 2017). In this study, we use *sentiment analysis* in *TextBlob* to measure the subjectivity. The procedure of sentiment analysis is as followed. After receiving a text sample, the package will process each word in the sample. As each word is processed, *TextBlob* searches its library for a match, and if a match occurs, the appropriate value for that word is incremented. At the end of this procedure, a final score (between 0 and 1) is assigned to each text sample, representing the extent of text objectivity and subjectivity where 0 is very objective and 1 is very subjective. Sentence examples with different levels of subjectivity are presented in Table 11. After getting the subjectivity score, we process it to a dummy variable by using the mean value of this score to separate the data to two parts with 1 indicating the score is above the average value and 0 indicating the score is below the average value. We conduct this process to mimic the experiment environment that we have only two conditions of review content (concrete vs. abstract). The average value of subjectivity is around 0.5, suggesting that the distribution of subjective/objective reviews is rather symmetrical.

Table 11 Examples of different score of review subjectivity

Subjectivity score	Example
0	They overprice the caskets so you can get the price reduced if you haggle.
0.5	Food is OK, atmosphere is meh. This might be a good place to watch the game.
1	Absolutely charming and delicious. I go here once or twice a week.

Another moderating variable of interest, skewness, is a representation of distribution asymmetry. It measures the relative position of mode and mean as well as the distribution shape, which is a more intuitive indicator. Figure 17 is an example of the three types of skewness. The one on the left is negative skewness, in which the reviewer is more likely to give positive ratings. The middle one is neutral skewness, which has an almost normal rating distribution. The one on the right represent positive skewness, indicating that the reviewer is likely to give negative ratings. All three distributions have similar means (around 3). As stated above, for potential customers who observe the distributions, they could infer the reviewers' rating habit. That is, if a reviewer who usually posts negative reviews (i.e., the right example) writes a negatively deviant review, potential customers may form some perceptions, such as "This reviewer follows his/her own habit. This review cannot reflect his/her own experience. This review has low quality!" However, if a reviewer who usually posts positive reviews (i.e., the left example) writes a negatively deviant review, potential customers may form other perceptions, such as “This reviewer breaks his/her own habit. This review must be a reflection of his/her own experience and is quite valuable!” Thus, skewness is an effective indicator of reviewer rating habit and accords with the operation of our first study. The calculation of the skewness of reviewers’ past rating distribution follows Fang et al. (2016)’s study. The average value of skewness is around 0.29, suggesting that the rating distribution is relatively balanced.

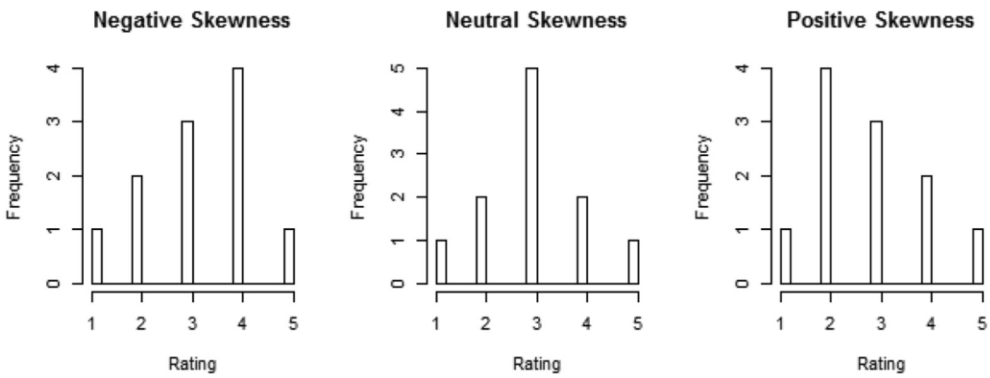


Figure 17 The example of different types of distribution skewness

Our analysis controls a series of variables from three levels (review level, reviewer level and restaurant level), including review length, review word length, reviewer join month, restaurant features. These variables have been used to study review perception/ in prior studies (Zhang et al., 2016; Fang et al., 2016; Yin et al., 2014).

On review level, we control review length and word length. Review length is operationalized as the number of words in a review. Longer reviews can provide more information and may influence customers’ perceptions. Review word length is operationalized as the average word length in a review. As a reflection of word/review complexity, word length can influence customers’ understanding toward a review. As the results of descriptive analysis, the average value of review length is 137 and that of review word length is 14.

On reviewer level, we control reviewer join month. Join month is operationalized as the number of months that a reviewer joins Yelp. We use September 2012 as the due date to track back the join month. Longer history of a reviewer may suggest a reviewer’s high credibility and influence customers’ perception toward his/her review. The average value of join month is 48 and the standard deviation of join month is 20, suggesting the join dates of reviewers have diversity.

On restaurant level, we control restaurant services including delivery and the device for wheelchair. These two variables may reflect a restaurant's attitude toward its customers and may influence its tendency to manipulate reviews, thus having an effect on customer perceptions toward its customer opinions. Delivery and wheelchair are dummy variables, with 1 indicating the existence of the services and 0 indicating the lacking of the services. The average value of delivery is 0.12, indicating that only a small ratio of restaurants provides this service; the average value of wheelchair is 0.67, indicating most restaurants try to provide convenience to disabled persons.

Table 12 Variable definitions

Variable	Definition
Manipulation	A dummy variable, representing whether a review is labelled as filtered review with 1 representing yes and 0 representing no.
Deviation	A dummy variable, representing whether rating for an individual review is negatively deviant from average rating with 1 representing the deviation while 0 representing no deviation.
Subjectivity	A dummy variable, representing whether the percentage of subjective words in a review is large with 1 representing a relatively large ratio of subjective words while 0 representing a relatively small ratio of subjective words. This variable is used to indicate the review content subjectivity.
Skewness	The skewness of rating distribution based on reviewers' past ratings, with larger skewness representing positively skewed rating distribution.
RevLength	Review length, represented by word number in a review.
WordLength	Word complexity, represented by the average word length in a review.
JoinMonth	How many months a reviewer joins Yelp.
Delivery	Whether a restaurant has delivery service with 1 representing yes while 0 representing no.
Wheelchair	Whether a restaurant has wheelchair service with 1 representing yes while 0 representing no.

The operationalization of all variables is summarized in Table 12. We also conduct the correlational analysis to rule out the multicollinearity. As Table 13 shown, the coefficients between variables are all lower than 0.8, indicating multicollinearity is not a serious problem in our research.

Table 13 Correlational analysis

	1	2	3	4	5	6	7	8	9
1.Fake	1								
2.Deviation	0.049	1							
3.FeelWord	-0.021	-0.0124	1						
4.Skewness	0.1244	0.0064	-0.009	1					
5.RevLength	0.0163	0.0072	-0.1368	0.0144	1				
6.WordLength	0.0045	0.001	-0.0891	0.0074	0.2367	1			
7.JoinMonth	0.0041	0.009	-0.0307	0.0388	0.0131	0.0528	1		
8.Delivery	0.0091	-0.006	0.011	0.0019	-0.0053	-0.0121	-0.0384	1	
9.Wheelchair	0.0067	-0.0456	0.0199	-0.0197	-0.0019	-0.014	0.0165	0.1124	1

2.6.3 Data analysis and results

Analysis is performed following the approach of prior studies (Bapna et al., 2011; Sinha & May, 2004), by using Logistic regression (one type of Logit regression) to analyse all reviews meeting the criteria described above (N = 8,487). We deem this approach appropriate because the dependent variable was a dummy variable. Table 14 presents the results of our empirical analysis. The results in all analysis indicate a good fit, with highly significant likelihood ($p < 0.001$) and sufficiently large values of Log likelihood.

In the first model of the regressions, we only include the independent variable (review deviation); in the second model, we also include a set of control variables with the independent variable; in the third model, we further add the moderating variables (review subjectivity and rating skewness); in the last model, we include all variables including the interactive items.

As the results shown, the coefficients of review deviation in all models are positive and significant ($\beta = 0.355$, $p<0.01$; $\beta = 0.358$, $p<0.01$; $\beta = 0.358$, $p<0.01$; $\beta = 0.355$, $p<0.01$). That is, reviews with negatively deviant rating from average rating are more likely to be manipulated. The results are in the expected direction, suggesting that H1 is supported. The coefficient of the interactive item between review deviation and review subjectivity is positive and significant ($\beta = 0.308$, $p<0.1$). In other words, as review content becomes more subjective, deviant rating has a larger chance to be manipulated, which is consistent with the hypothesis expectation. Thus, H2 is supported. In addition, the coefficient of the interactive item between review deviation and rating skewness is positive and significant ($\beta = 0.245$, $p<0.01$). A larger value of skewness suggests negative rating habit (positively skewed distribution in first study) and a lower value of skewness suggests a positive rating habit (negatively skewed distribution in first study). That is, a review with negatively deviant rating is likely to be a manipulated one if its reviewer has negative rating habit in the past. In contrast, a review with negatively deviant rating is less likely to be a manipulated one if its reviewer has positive rating habit in the past. Thus, H3 is supported.

Logistic regression results also involve some control variables. However, the control variables show non-significant effect on review manipulation. Simple indicators of reviewers and restaurants such as review length and reviewer join month cannot predict the review credibility. Thus, in Yelp, these indicators have little associations with the judgement of review manipulation.

Table 14 Logistic analysis on review manipulation

VARIABLES	Model 1 Manipulation	Model 2 Manipulation	Model 3 Manipulation	Model 4 Manipulation
Deviation	0.355***	0.358***	0.358***	0.355***
	(0.079)	(0.079)	(0.080)	(0.119)
Deviation* Subjectivity				0.308*
				(0.159)
Deviation*Skewness				0.245***
				(0.069)
Subjectivity			-0.126*	-0.216**
			(0.074)	(0.088)
Skewness			0.359***	0.189***
			(0.032)	(0.057)
RevWordNum		0.013	0.009	0.011
		(0.049)	(0.050)	(0.050)
WordLength		0.083	0.059	0.052
		(0.112)	(0.114)	(0.114)
JoinMonth		0.007	-0.020	-0.021
		(0.067)	(0.068)	(0.068)
Delivery		0.086	0.082	0.084
		(0.108)	(0.109)	(0.109)
Wheelchair		0.057	0.077	0.072
		(0.078)	(0.078)	(0.079)
Constant	-2.421***	-2.784***	-2.722***	-2.717***
	(0.066)	(0.364)	(0.373)	(0.379)
Observations	8,487	8,487	8,487	8,487
Chi2	20.96	23.42	156.78	173.07
R^2	0.004	0.004	0.028	0.031

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

2.7 Discussion and conclusion

Based on the increasing number of fake reviews appearing on online platforms (Luca & Zervas, 2016), focusing on customers' perceptions of online reviews concerning the possibility of fake reviews can offer important insights for researchers and practitioners. In this study, we propose a new concept—perceived review manipulation—which denotes the extent to which an individual perceives a review is manipulated. This concept derives from but differs from perceived review credibility by focusing on the aspect of review untrustworthiness.

Reviews with deviant ratings are distinct from other reviews and are more likely to attract customer attention; thus, researchers and practitioners have begun to investigate their influence on customer perceptions, including perceptions of review helpfulness. Because previous research has not clarified the influence of deviant ratings on perceptions of review fakery, this study seeks to fill that research gap by exploring how deviant ratings affect perceived review manipulation. Prior studies have examined the effect of reviews with deviant ratings on perceived review usefulness but have generated contradictory findings (Gao et al., 2017; Qiu et al., 2012). Thus, we seek to clarify this effect by proposing that perceived review manipulation can work as a novel mechanism articulating the influence of deviant ratings on customer perceptions related to purchase decisions (i.e., perceived review helpfulness and adoption).

This study also investigates how deviant ratings, review content, and reviewer information jointly affect customers' perceptions of review manipulation. Existing studies mostly focus on how the quantitative metrics of online reviews such as review valence and variance moderate the influence of deviant ratings on review perceptions (Yin et al., 2016; Qiu et al., 2012) and do not clarify how review-related qualitative factors such as review content and reviewer information contribute. Review content concreteness is a determinant of review perception independently or when combined with other factors such as review timing (Huang et al., 2018). Review distribution based on reviewers' past ratings can reflect their rating history and habits, which can also influence how their reviews are perceived by other potential customers (Fang et al., 2016). Thus, identifying the moderating role of review content concreteness and reviewer rating distribution can significantly improve the understanding of the effect of rating deviation on perceived review manipulation.

Using data from Amazon Mechanical Turk, this study designed two online randomized experiments to (1) investigate the effect of rating deviation on perceived review manipulation and identify the boundary conditions concerning review content and reviewer features, and (2) explore how rating deviation and perceived review manipulation work on the perceptual indicator of customer purchase decisions (perceived review helpfulness and adoption). The findings suggest a positive impact of rating deviation on perceived review manipulation. Reviews with deviant ratings are less likely to be perceived as manipulated if their content is concrete rather than abstract. In addition, deviant reviews written by reviewers with a negatively skewed past rating distribution are less likely to be perceived as manipulated, whereas deviant reviews written by reviewers with positively skewed past rating distributions are more likely to be perceived as manipulated. The results also suggest a negative influence of rating deviation on perceived review helpfulness and adoption since a deviant review is more likely to be perceived as manipulated. The findings imply that perceived review manipulation is a stable mechanism underlying the effect of rating deviation on perceived review helpfulness and adoption and that factors such as review content and reviewer information can moderate the impact of rating deviation.

As Yelp proposes a relatively objective evaluation on review manipulation, this study takes advantage of this particular design and employs data collected by a prior study to validate the results found in the experiment. Specifically, the empirical study using Yelp data is to (1) explore the effect of rating deviation on review manipulation; (2) investigate the moderating role of review content subjectivity (vs. objectivity) and the skewness of reviewer rating distribution on review manipulation. The results are consistent with those found in the experiments. Specifically, a review with deviant rating is likely to be manipulated. If this review has subjective content, the probability that it is manipulated becomes further bigger. If this review is written by a reviewer with larger skewness (negative rating habit), it is more likely to be manipulated. If this review is written by a reviewer with smaller skewness (positive rating habit), it is less likely to be manipulated. The findings suggest that our hypotheses are still robust in real world setting.

2.7.1 Theoretical implication

This study makes significant contributions to the literature in three ways. First, this study contributes to the online review literature. Prior studies mainly explore the effect of rating deviation on perceived review helpfulness (Yin et al., 2016; Qiu et al., 2012), while little attention has been devoted to the impact of deviation on review adoption, which directly impacts customer purchasing decisions (Cheung et al., 2008). The present study, suggests that review deviation and perceived review manipulation can impact review adoption, extending existing studies for another dimension that could be impacted by rating deviation. In addition, rating deviation in prior studies involves positive and negative average ratings, as well as individual ratings (Yin et al., 2016). This study fixes the average rating as positive and manipulates rating deviation by adjusting only the individual rating. Compared to previous studies considering both positive and negative average rating, fixing the valence of average rating can be helpful to isolate the impact of rating deviation and explain the contradictory results presented in previous research (Gao et al., 2017; Shen et al., 2015; Danescu-Niculescu-Mizil et al., 2009).

Second, this study extends our understanding of the review manipulation literature by proposing perceived review manipulation as a new dimension as customer perception. While a few prior studies have begun to pay attention to review manipulation/fake reviews on how review manipulation can damage firm profits (Anderson & Simester, 2014; Hu et al., 2012), little attention has been paid to how customers perceive reviews as manipulated. This study is one of the very first studies to focus on this novel research direction by proposing the concept of perceived review manipulation and developing items to measure this construct based on previous research (Zhang et al., 2016; Hu et al., 2012; Mayzlin, 2006). This study also complements prior studies by identifying factors that contribute to perceived review manipulation. In addition, this study takes advantage of Yelp's filtering algorithm and employs online secondary data to validate the results concerning to review manipulation. Compared to prior studies on fake review topic, our research combines controlled experiment data and online field data to give more detailed exploration of the impact of review rating, review content and reviewer rating distribution. This mixed-method can also help us understand whether customers' perception on review manipulation is consistent with the real-world evaluation about review manipulation. Third, this study provides empirical evidence through investigating the joint effect of review content, reviewer rating distribution, and rating deviation, thus extending previous knowledge in recognizing the boundary conditions of the effect of rating deviation and filling the gap in the literature concerning review content concreteness and reviewer rating distribution. In prior studies, the direct effect of review content concreteness has been identified (Shin et al., 2019; Huang et al., 2018; Wang et al., 2015); but the moderating role of concreteness has not been examined. Our findings suggest that concrete review content reduces the effect of rating deviation on perceived review manipulation. Furthermore, since little previous research has focused on reviewer rating distributions (Fang et al., 2016), our findings that link reviewer rating distributions to potential reviewer rating habits contribute to this literature.

2.7.2 Practical implications

This study also has several practical implications. First, in browsing online reviews, some consumers tend to trust individual ratings that correspond to average ratings, whereas others may find deviant or emotional opinions to be more trustworthy. Our results suggest that customers should evaluate review content and reviewer rating trends rather than simply relying on rating deviation/consistency since such information can also indicate review quality and reviewer credibility, both of which are helpful for making informed purchase decisions.

Second, our results indicate that reviews with deviant ratings tend to be perceived as manipulated and less helpful and are less likely to be adopted. However, such biases may be largely unsupported. Thus, we recommend that platform managers take action against such biases, in order to preserve the influence of high-quality reviews that may otherwise be overlooked, leading to potential negative impacts for the platform. For example, platforms should actively encourage reviewers to write more concrete reviews with detailed descriptions, perhaps by offering badge-related incentives. We recommend that platforms highlight rating distributions based on reviewers' past ratings because they are useful for evaluating review manipulation and can help alleviate customer bias toward deviant reviews.

Third, since reviews with deviant ratings are less likely to be adopted, this study suggests that reviewers, especially those who give varied opinions, write concrete reviews with specific descriptions of their experience. Reviewers should also become conscious of their reviewing habits and seek to devote effort to each review they write. Otherwise, their opinions, especially when they differ from others, may be perceived as manipulated.

Finally, secondary data from Yelp indicate that reviews with deviant ratings are indeed likely to be manipulated and review content subjectivity as well as reviewer rating distribution can work in real world setting. We thus suggest the platforms to consider the two aspects (review content subjectivity and reviewer rating distribution) when they optimize their filtering algorithm in the future. In this way, they can provide customers more accurate guidance on filtered reviews.

2.7.3 Limitations

This study also has several limitations, which, however, provide important avenues for future research. First, as discussed in prior studies, other aspects of review content such as review argumentation (density and diversity), review sentiment, and reviewer-related characteristics such as reviewer identity disclosures can also affect perceived review helpfulness (Willemsen et al., 2011; Felbermayr & Nanopoulos, 2016). To isolate the effect of content concreteness and reviewer rating distributions, this study disregards issues potentially related to other factors. Future studies should investigate the moderating role of these factors to validate and extend our results. In addition, this study only focuses on negatively rating deviation to better extract its effect on perceived review manipulation. Although this kind of deviation is more likely to happen in real life (most restaurants have positive average rating), future studies are encouraged to extend the research to the context with positively deviation rating.

Second, while this study recruited participants from North America to improve consistency, an interesting avenue of research would be the exploration of different cultural influences. For instance, customers from more individualistic cultures may be more willing to accept divergent opinions, as compared to those from collectivist cultures (Liu, Rob, & Xu, 2018).

Third, the findings in this study exclusively used an experience product (i.e., hotel). Compared to experience products, purchase considerations for search products such as cameras, books, etc., are likely to be more objective and less emotional, suggesting that reviews with deviant ratings may be more likely to be perceived as manipulated rather than as a reflection of the reviewer’s own feelings. Thus, future research could investigate whether customers are more likely to consider deviant ratings as manipulated in the context of search products.

Fourth, this study explores the moderating role of reviewer rating distribution, which is a practical design in Yelp. To my best knowledge, this practical design is unique in Yelp at present, which may restrain the practical implications of this study. However, as Yelp is one of the most popular platforms to provide guidance for customers to choose products like restaurants, the effect of its features still deserves investigation.

Chapter 3

Arousal level, personal involvement and individual ability to distinguish fake news

3.1 Introduction

The breakout and spread of COVID-19 have influenced individuals from all over the world. What spreads faster than the virus is the fake news about COVID-19. Between January 22nd (the date that Wuhan China was locked down by Chinese authorities) and March 14th, as the observatory showed that around 275,000 Twitter accounts posted 1.7 million links to unreliable news about the virus (Hernandez, 2020[1]). The deliberate creation and spread of fake news may cause public panic and even social instability. This is not the first time that fake news misleads public during important events. In the 2016 with US presidential election, the prevalence of fake news largely influences the election results and receives global attention (Allcott & Gentzkow, 2017). Because of the worldwide coverage of COVID-19, fake news toward COVID-19 appeals attention from academia and industry.

Fake news is real-time, making it hard for platforms to detect and screen. Thus, individuals often need to distinguish fake news by themselves. One important question is whether individuals has the ability to distinguish fake news from real news. Unfortunately, individuals' ability on fake news detection is limited. According to a research by Pew Research Center, 23% of Americans had shared a made-up news story with or without conscious (Barthel, 2016). Individuals' low ability to discern fake news may come from confirmation bias (Kim et al., 2019), illusory truth effect (Pennycook et al., 2018) and personal features (Verma et al., 2018). As fake news toward COVID-19 may cause panic and influence the fight for the virus, what factors can determine individuals' ability to judge fake news toward COVID-19 deserves further attention.

Epidemic event such as COVID-19 is terrible and may cause death, thus information concerning to this topic may elicit individuals' emotional responses. As emotions may evoke different physiological arousal and then influence individual rationality (Smith & Ellsworth 1985), we intend to focus on arousal level triggered by news content as one influential factor

[1] https://www.msn.com/en-us/money/other/fake-news-about-covid-19-is-spreading-faster-than-the-virus/ar-BB11N2XY

on individuals’ ability to judge fake news toward COVID-19. Epidemic event poses impact to almost everybody and individuals are forced to be influenced from the event themselves. Thus, individuals’ personal involvement toward the epidemic news is another focused factor in this study. We conducted a pilot study and a formal study by recruiting participants from four Chinese universities to investigate the role of arousal level and personal involvement on individuals’ ability to distinguish fake news. The findings suggest that individuals initially have ability to discern fake news, but they tend to trust and share fake news if news content can trigger a high arousal level. In addition, with a high personal involvement toward the news of COVID-19, individuals are more likely to trust fake news with content triggering a high arousal level. The findings contribute to existing studies by introducing arousal level and personal involvement in the context of fake news and provide guidance for news platforms or social media platforms about what news should be governed eagerly.

3.2 Past studies and theoretical foundation

3.2.1 Fake news

The term “fake news” has a long history and is a form of disinformation that mimics the look and feel of news (Gelfert, 2018). It has been used to describe political satire, news parody, and even misleading advertising (Tandoc et al., 2017). Recent studies use fake news to refer to all kinds of false stories or news spread on the Internet such as social media platforms (Zhang & Ghorbani, 2020). In other words, fake news can represent news articles that are intentionally and verifiably false, and could mislead readers (Allcott & Gentzkow, 2017).

With the absence of verification mechanism, everyone can easily write fake news on the Internet (Ahmed, 2017). Thus, large volume of fake news exists on social media and causes serious outcomes such as 2016 election results in the United States (Allcott & Gentzkow 2017). The influence of fake news is not confined to only elections but also brings fateful consequences to the other areas including IT industry, and financial markets and whole society that everyone may live in a cyber-environment with the crisis of trust (Zhang & Ghorbani, 2020). For example, in the long run, uncontrolled fake news can undermine public trust toward other information sources even if the sources are credible (Lazer et al., 2018)

Prior studies have investigated computational tools that can identify fake news on social media (Shu et al. 2017). Fake news is related to four major components (fake news creator, target victim, fake news content, and fake news social context), thus its detection approaches can be divided as creator and user analysis, news content analysis and social context analysis (Zhang & Ghorbani, 2020). Creator and user analysis mainly refer to user profiling analysis such as the geographic locations of the account (Zhang et al., 2020), temporal and posting behavior analysis such as the average time between two news (Chu et al., 2012), and sentiment-related analysis such as the usage of happiness score to indicate the original creator's emotion (Dodds et al., 2015). News content analysis contains several aspects including linguistic and semantic-based analysis and style-based analysis. In linguistic and semantic-based analysis, the main point is to observe language formats and discover writing patterns using natural language processing methods such as long-short-term memory (LSTM) neural network (Ajao et al., 2018). In style-based analysis, text complexity and readability can be used to identify the authenticity of the news (Horne & Adali, 2017). Social context analysis is realized by the network analysis of the news creator (Ruchansky et al., 2017) and distribution pattern analysis (Jin et al., 2016).

Fake news creators tend to be short-lived, causing a high velocity of fake news appearance (Allcott & Gentzkow, 2017). The real-time nature of fake news makes it hard for social media platforms to identify and tag fake news on time. Thus, it's important to investigate the detection of fake news from the perspective of news readers. That is, who are more likely to trust/share fake news and what platform designs are helpful for readers to evaluate fake news? Several studies have already focused on the factors that can influence readers' perception on fake news including readers' own characteristics and platform designs (Kim et al., 2019; Kim & Dennis, 2019; Colliander, 2019; Verma et al., 2018; Pennycook et al., 2018). For example, changing presentation format to highlight news source can make users more skeptical of all articles, regardless of the source's credibility (Kim & Dennis, 2019). In addition, the presentation of critical comments from others toward a piece of fake news significantly influences individuals' attitude towards the news, as well as their intentions to comment and share the news (Colliander, 2019).

3.2.2 Emotional arousal

Several models classify fundamental dimensions of individuals' emotional experience (Eaton & Funder, 2001; Carstensen et al., 2000; Lanzetta et al., 1976). Among the dimensions of emotional experience, arousal is of great importance (Kron et al., 2015). It is a feeling of intensity and describes the extent to which the automatic nervous system of an individual is activated by an experience, varying from drowsiness to excitement (Heilman, 1997; Mehrabian & Russell 1974). That is, high arousal or activation is characterized by activity and low arousal or deactivation is characterized by relaxation. The measurements of arousal involve several dimensions of emotions such as 'passive' vs. 'active' and 'mellow' vs. 'fired up' (Berger; 2011).

Prior studies have shown that activated arousal can influence individual behaviors or perceptions in several areas including organizations (Pazzaglia et al., 2012; Brooks & Schweitzer 2011), sociology (Swann et al., 2010; Gaertner & Dovidio, 1977) and review evaluations (Yin et al., 2017). For example, when witnessing an emergency situation, a bystander's responsiveness to help others becomes much quicker if the bystander experiences more arousal for the emergency (Gaertner & Dovidio, 1977). Arousal level have similar effects on social transmission and sharing possibility in online content such as advertisements and political articles (Rubenking, 2019; Guadagno et al., 2013; Okdie et al., 2013). Specifically, individuals report a greater intention to spread a video which can trigger their stronger affective responses (Guadagno et al., 2013).

Arousal can be induced through either physical exercise (i.e., running time) or emotional material (i.e., advertisement content) (White et al., 1981). In online context, physical contact is barely happen. Thus, emotional material becomes the most common way to induce individual arousal. For example, with the usage of different musical pieces in advertisements named "Eine Kleine Nachtmusik: Allegro" by Mozart and "Whatever We Image" by David Foster, individuals will be manipulated with different self-report arousal levels (Gorn et al., 2001). In addition, online content that evokes emotions such as anger and anxiety is more arousal than that evokes emotions including sadness (Berger & Milkman, 2012).

3.2.3 Personal involvement

Involvement refers to individual's perceived relevance of the object based on inherent needs, values, and interests (Zaichkowsky, 1986). It is considered an important psychological construct that influences the allocation of cognitive resources to evaluation of an object, decision, or action (Mitchell, 1979). Houston & Rothschild (1978) classified involvement into three dimensions including situation involvement, enduring involvement, and response involvement. Situation involvement refers to an individual's temporary feelings of involvement that accompanies a particular situation, focusing on an individual's non-personal factors (Huang et al., 2010; Richins et al.,1992). Houston & Rothschild (1978) proposed two categories of stimuli of consumer's situation involvement: product-related and social psychological stimuli, which determined the level of situation involvement. Enduring involvement represents the general, long-run concern with objects or things that individuals bring to a situation, which emphasizes personal characteristics (Houston & Rothschild, 1978; Huang et al., 2010; Richins et al., 1992). Situation involvement and enduring involvement jointly influence response involvement, which signifies an individual's decision-making based on the complexity of cognitive and other processes at various stages (Houston & Rothschild, 1978; Huang et al., 2010; Richins et al., 1992).

Involvement condition is one of the crucial determinants of individual behaviour (Britwum & Yiannaka, 2019; Guo et al., 2019). Scholars in the field of marketing have proposed numbers of concepts and theories related to involvement in consumer decision-making processes, such as product involvement (Cowan & Ketron, 2019; Harun & Prybutok, 2020; Hong, 2015), consumer involvement (Cruz et al., 2017; Yoon & Zhang, 2018), and brand involvement (Bian & Haque, 2020). In particular, scholars have certificated that consumers' product/brand involvement is a good predictor of consumer purchase intention (Hong, 2015; McClure & Seock, 2020), service/product satisfaction (Gohary et al., 2016), consumer loyalty (Chen & Tsai, 2008; Harun & Prybutok, 2020; Wang et al., 2006; Wu & Hsiao, 2017), brand attitudes (Bian & Haque, 2020; Spielmann & Richard, 2013), and consumer engagement (Liu & Jo, 2020). Drawing on previous studies, we believe that personal involvement plays an important role in individual decision-making process and profoundly affects individual intentions and behaviours.

The COVID-19 epidemic is an important global event, posing impacts on individuals' daily life because of its global spread. Therefore, a large number of individuals are very actively concerned about the epidemic news. We thus believe that epidemic event involvement will profoundly affect the individual's cognition and judgment toward related news.

3.3 Hypotheses development

The concept of rationality suggests that individuals make judgments such as criticizing someone else's thoughts or defending their own thoughts with rationality all the time (Doyle, 1992). In general, individuals have limited rationality; thus, they may be able to make judgements of information in life. While this rationality may be influenced by particular reasons such as individuals' strong emotions or individuals' high concern toward an event. I propose the following hypotheses under the concept of rationality.

3.3.1 Reader judgement on fake news

Individuals' trust and willingness to share are two main focus when talking about disinformation (i.e., rumor and fake news) (Visentin et al., 2019; Kim & Dennis, 2019). Trusting and sharing intentions have been sometimes investigated in one study as these two aspects are different from each other even with some similar meanings (Kim et al., 2019; Colliander, 2019; Chua & Banerjee, 2018). That is, individual' trusting perception toward information does not always lead to their sharing behaviour (Seifert et al., 2017) and individuals may still choose to share information because of some motivations such as social norms and altruism to other users even though they distrust the information (Burtch et al., 2017; Ma & Chan, 2014; Kankanhalli et al., 2005). Thus, this study examines both trusting and sharing intention to share to better understand individuals' perception on fake news under different situations.

Several distinctions exist between real news and fake news as the former one is realistic and the latter one is pretended to be realistic. From the content perspective, articles in fake news tend to be shorter and contain more repetitive language and fewer punctuation; titles in fake news are more likely to be longer, use fewer stop words and nouns but more proper nouns (Horne & Adali, 2017). Generally speaking, information in real news may seem more engaging and has little room for individuals to discount (Balmas, 2014; Busselle et al., 2000). Based on the concept of rationality (Doyle, 1992), when confronting news with uncertain sources on social media, individuals are likely to evaluate the content in the news and make rational judgements to discern fake news and real news. For example, on Seeking Alpha, investors in general are able to discern fake news as they can discount biased information even though fake news attract their significant attention (Clarke et al., 2019). In addition, social media platforms such as Facebook have started fact-checking mechanism to encourage users to indicate suspicious news for platforms to double check (Meinert et al., 2018). With this particular design, users may sometimes keep skeptical about the news credibility and are able to detect fake news by themselves (i.e., seek verification through personal channels) (Tandoc et al., 2018). Then they may form different trusting and sharing intention toward fake news and real news. Thus, we propose the following hypotheses:

H1a: In normal situations, individuals are more likely to trust real news than fake news.

H1b: In normal situations, individuals have higher sharing intention to real news than to fake news.

3.3.2 The role of emotional arousal

Heighted arousal can lead to low-quality judgements such as polarized/similar evaluations (Gorn et al., 2001) and poor task performance (Valiente et al., 2012). Two reasons can explain this negative effect. On one hand, the level of arousal is associated with the range of attention (Gellatly & Meyer, 1992; Easterbrook, 1959). When individuals' arousal level in a task situation increases, the range of attention decreases that individuals have to pay attention to specific parts of the task and may ignore some relevant or important information, thus the performance may be restrained (Humphreys & Revelle, 1984). One the other hand, a high level of arousal results in an emotion-driven thinking under which individuals tend to automatically process and utilize emotion-related words even with the existence of other sufficient information (Gendron et al., 2012; Gernsbacher et al., 1998; Angrilli et al., 1997). In other words, high arousal level leads individual to process less complex information such as peripheral cues (i.e., background music and source attractiveness) (Sanbonmatsu & Kardes, 1988; Chaiken, 1980), which may disrupt the traditional route and reduce individuals' capacity of information processing especially when individuals deal with complex tasks (Zajonc, 1965). In this way, individuals' information comprehension and task performance may reduce.

Extending to our context, when social media news with specific content makes individuals become high aroused, they may form a narrow range of attention and a reduced ability to process news information (Easterbrook, 1959). Under this situation, they may allocate their limited attention to irrelevant aspects of the news such as the character style rather than important cues such as the news content (Chaiken, 1980). That is, individuals' ability to integrate relevant aspects of news during cognitive process may be reduced and their evaluation toward fake news and real news may become biased and incorrect (Ellis & Ashbrook, 1988). In this way, individuals may express high trusting perception and sharing intention toward fake news than usual. Thus, we propose the following hypotheses:

H2a: When news contents trigger high arousal of the individuals, they show similar trusting toward fake news and real news. When news contents trigger low arousal of the individuals, they tend to trust real news rather than fake news.

H2b: When news contents trigger high arousal of the individuals, they show similar sharing intention toward fake news and real news. When news contents trigger low arousal of the individuals, they tend to form higher sharing intention toward real news rather than fake news.

3.3.3 The role of personal involvement

High level of involvement is easier to trigger individuals' specific intentions and behaviours toward specific events. For example, in marketing context, the level of product involvement is usually associated with consumers' purchase behaviours (Hong, 2015; McClure & Seock, 2020) and high level of the consumer involvement is often be regarded as a positive predictor of consumer loyalty (Harun & Prybutok, 2020; Wang et al., 2006; Wu & Hsiao, 2017). In the context of political election, political involvement is sometimes considered as a symptom of political participation and police support (Gohary et al., 2016; Guo et al., 2019). According to the definition of the situation involvement proposed by Houston and Rothschild (1978), epidemic event involvement refers to individuals' perception of epidemic event based on their inherent needs, values and interests. That is, a high level of epidemic event involvement may drive individuals' interest and attention to related information (i.e., news).

Natural disasters, public health events, and other major events often lead to negative public sentiment such as fear, anxiety, disappointment, nervousness, sadness, etc (Cheliotis, 2020; Li et al., 2020; Seltzer et al., 2017). These negative emotions will lead to varieties of adverse effects on individuals, for instance, individuals' reduced cognitive level and decision quality (Lerner & Keltner, 2000, 2001). Extending to our context, when individuals are highly involved in the COVID-19 epidemic event, they may pay more attention to the related information (i.e., news). As most news of epidemic event contain negative information such as the rising number of infections and deaths, depressed economy, or even riots, negative emotions may dominate public sentiment, resulting in individuals' emotion-driven thinking and a decline of individual's judgment ability toward the news. If the individuals are in high arousal level at the same time, their lower ability to process the news content may be further reduced and it may be more difficult for them to distinguish between true and false news (Ellis & Ashbrook, 1988), which in turn can lead individuals to trust/share fake news and real news without difference. However, if individuals are not highly involved in the epidemic events, their ability to discern real news and fake news may be only influenced by their arousal level. Thus, we propose the following hypotheses:

H3a: For individuals who are highly involved in the epidemic event, their ability to distinguish fake news and real news becomes lower especially when the content triggers high arousal rather than low arousal. For individuals who are not highly involved in the epidemic event, their ability to distinguish fake news and real news is similar to general situations.

H3b: For individuals who are highly involved in the epidemic event, their tendency to share real news rather than fake news becomes lower especially when the content triggers high arousal rather than low arousal. For individuals who are not highly involved in the epidemic event, their intention to share real news and fake news is similar to general situations.

The research framework is shown in Figure 18.

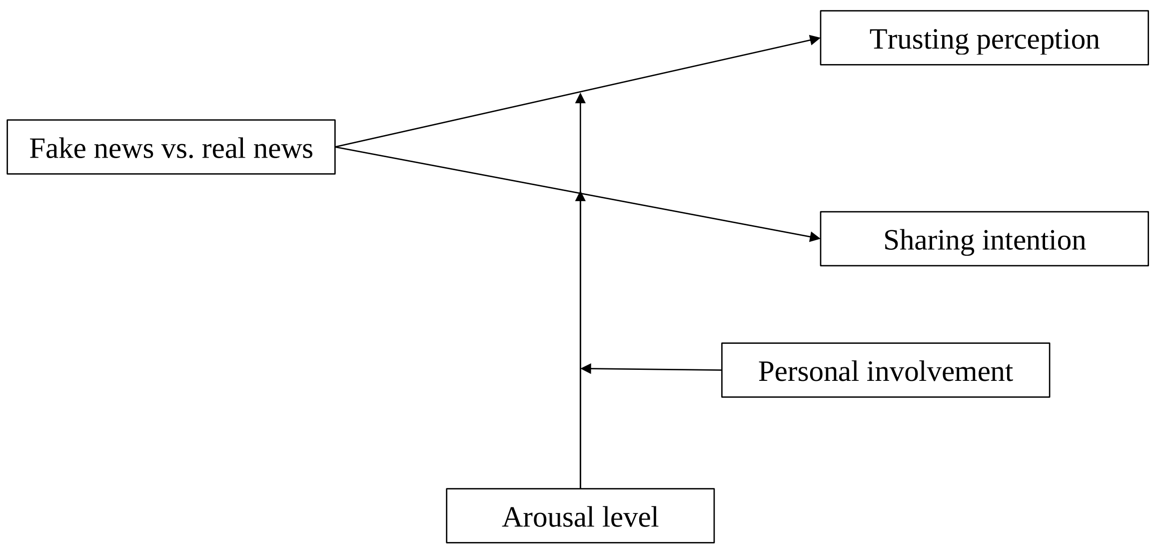


Figure 18 Research framework

3.4 Pilot study

Pilot study was intended to validate the appropriateness of the materials (news) in the formal experiment. That is, whether the news with different content can effectively trigger different individual arousal levels. To that end, we used a 2*2 between-subjects experimental design. One group of participants (group 1) was subjected to fake news with angry content. The second group of participants (group 2) was subjected to fake news with sad content. The third group of participants (group 3) was subjected to real news with angry content. The last group of participants (group 4) was subjected to real news with sad content. We used sad content and angry content to trigger different arousal levels following prior studies (Berger & Milkman, 2012; Berger, 2011).

3.4.1 Experiment procedure

This study recruited participants from student pools of a Hong Kong university. For participants who successfully registered and finished the experiment, a credit will be rewarded. At last, we recruited 100 participants with about 25 participants for each group to test the validity of news materials.

After the recruitment, participants would receive a link in which they were assigned to one of the four groups that were embedded in a survey tool (Qualtrics). Participants were first instructed to imagine that they were browsing news about COVID-19 and one piece of news showed up. After reading the news, participants were asked some questions referring to emotions and arousal level.

3.4.2 Manipulations and Measures

This study focused on individuals' arousal perception on different news content. To manipulate news to trigger different arousal level, a news title with a related picture was clearly presented to participants. Both fake news and real news were manipulated to trigger different arousal level. In the fake news with low arousal condition, a news that contains sad information and has been tagged as fake is presented. In the real news with low arousal condition, a news that contains sad information and has been proved as real is presented. In the fake news with high arousal condition, a news that contains angry information and has been tagged as fake is presented. In the real news with high arousal condition, a news that contains angry information and has been proved as real is presented.

Manipulation checks focus on participants' anger emotion after reading the news by asking participants to choose from 7-point scales about the question: "You are angry after reading the news.". Arousal level, the main variable in pilot study, was measured using three 7-point scales (passive—active, mellow—fired up, low energy—high energy) by one question adopted from Berger (2011): "How do you feel after reading the news?"

3.4.3 Results

We expected that angry emotion in the groups of fake news or real news with angry content will significantly exceed that in another two groups. ANOVA analysis revealed a significant effect of news with angry content on angry emotion ($F = 26.45$, $p < 0.01$), and the mean value of angry emotion is larger in the angry-content condition (mean = 5.277, S.D. = 0.182) than in the sad-content condition (mean = 3.863, S.D. = 0.204). The successful manipulation suggests that compared to news with sad content, news with angry content can trigger angry emotion of participants. The results were also presented in Figure 19. We also expected that arousal level in the groups of fake news or real news with angry content will significantly exceed that in another two groups. We also conducted ANOVA analysis and the results revealed a significant effect of news with angry content on arousal level ($F = 11.36$, $p < 0.01$). The mean value of arousal level is larger in the angry-content condition (mean = 3.950, S.D. = 0.119) than in the sad-content condition (mean = 3.333, S.D. = 0.140). The results were also presented in Figure 20. The results verified the appropriateness of using news with sad and angry content in formal study to trigger participants' arousal and test the hypotheses.

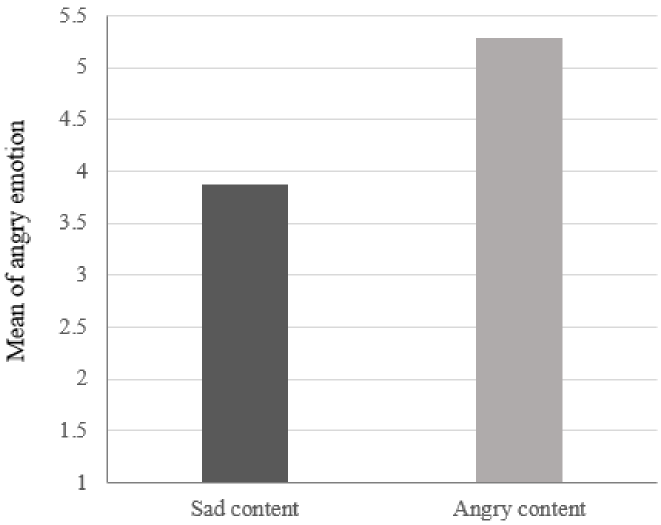


Figure 19 The effect of news content on angry emotion

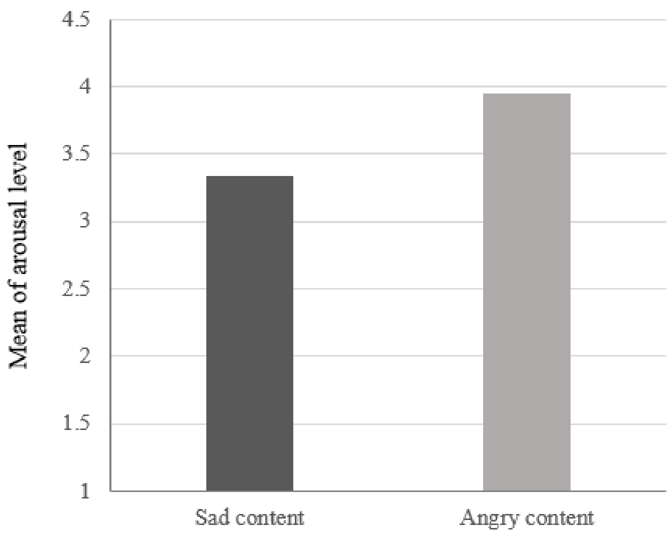


Figure 20 The effect of news content on arousal level

3.5 Formal study

Formal study was intended to test the hypotheses using news materials validated in pilot study. That is, whether individuals have the ability to distinguish fake news and real news in normal situation (H1) and whether individuals' discern ability toward fake and real news changes when news content can trigger different arousal level (H2). In addition, for individuals with different involvement level toward the epidemic, whether their ability to detect fake news shows any differences (H3). To test these hypotheses, we conducted a 2*2 between-subjects experimental design with fake news and real news combined with high-arousal content and low-arousal content. The experimental groups are the same as in pilot study. We translated the news content to recruit Chinese students as they are involved in the epidemic, which is helpful to test H3.

3.5.1 Experiment procedure

The formal study recruited students from three universities in Mainland China. After finishing the experiment, the reward was allocated to participants using a red envelope with random reward from RMB 1 to RMB 5. Similarly, after the recruitment, participants would receive a link in which they were assigned to one of the four groups that were embedded in a survey tool (WJX.cn). To rule out the possible influence of participants' past history of news reading, we first asked participants whether they have read or known the news before. Only participants who haven't read the news before can continue the experiment and receive the final reward. The other procedure was similar with that in pilot study that participants were asked to imagine that they were browsing news about COVID-19. After showing up the news, participants were asked some questions referring to emotions and arousal level, personal involvement, believability perception and sharing intention toward the news.

3.5.2 Manipulations and Measures

This study focused on the effect of arousal level of news content and personal involvement on trusting perception and sharing intention toward fake news and real news. Similar to pilot study, a news title with a related picture was presented to participants to manipulate news with different authenticity and arousal. The operationalization of news content is translated from that in pilot study (shown in Figure 21).

One moderator, arousal level, is measured using three questions translated and adapted from Berger (2011): (1) "阅读完这则新闻, 我感觉自己内心难以平静 (After reading this news, it's hard for me to calm down)," (2) "阅读完这则新闻, 我感觉自己思绪起伏不定 (After reading this news, I feel my feelings fluctuate)," (3) "阅读完这则新闻, 我感觉自己内心震颤 (After reading this news, I feel tremor inside my heart)". Another moderator, personal involvement, is measured using five questions that are adapted and translated from Zaichkowsky (1985): (1) "这则新闻引起了我的关注 (This news catches my interest)," (2) "这则新闻引起了我的注意力 (This news catches my attention)," (3) "我愿意持续关注这一新闻 (I am willing to pay attention to this news in the future)," (4) "这则新闻对我了解疫情信息很有帮助 (This news is helpful for me to understand epidemic information)," (5) "我这则新闻对我了解疫情信息很有价值 (This news is valuable for me to understand epidemic information)".

Trusting perception, one dependent variable, is measured using three questions translated and adapted from Kim & Dennis (2019): (1) "我感觉这则新闻是可信的 (I feel this news is believable)," (2) "我感觉这则新闻是真实的 (I feel this news is truthful)," (3) "我感觉这则新闻是可靠的 (I feel this news is credible)". Sharing intention, another dependent variable, is also measured by a question translated and adapted from Kim & Dennis (2019): "我愿意分享和转发这则新闻".

NEWS

生命不平等！意大利医院停止向 60 岁以上的老人提供呼吸机，将省下来的资源留给年轻人！

By Paul Thompson April 10, 2020 | 11:30am | Updated

Table 21a Fake news with angry content

NEWS

乌克兰卫生部长将 65 岁以上的老人称为尸体，并呼吁不要为了支撑这些老人而浪费金钱和资源！

By Kelly Smith April 10, 2020 | 11:30am | Updated

Table 21b Real news with angry content

NEWS

为了陪伴自己身患新冠肺炎的女朋友，一个男人脱下自己的防护服，最终二人双双死于新冠肺炎

By Paul Thompson April 10, 2020 | 11:30am | Updated

Table 21c Fake news with sad content

NEWS

美国一疗养院的老人因患新冠肺炎，去世前太过疼痛，多次向手机的语音助手求助

By Kelly Smith April 10, 2020 | 11:30am | Updated

Table 21d Real news with sad content

Figure 21 News content operationalization

3.5.3 Results

We first checked the arousal level of news with angry content and sad content again before testing the hypotheses. Similar to pilot study, the results suggested that arousal level triggered by news with angry content significantly exceeds that triggered by news with sad content (F = 3.68, $p < 0.1$). Thus, we used news with angry content as high-arousal condition and news with sad content as low-arousal condition.

H1 posits that individuals can discern fake news and real news in normal situation that they will show more trusting perception and higher sharing intention toward real news compared to fake news. ANOVA analysis revealed a significant effect of news fakery (vs. authenticity) on the trusting perception of news (F = 4.56, $p < 0.05$), and the mean value of trusting perception of news is larger in the real-news condition (mean = 3.738, S.D. = 0.073) than in fake-news condition (mean = 3.515, S.D. = 0.074). Thus, H1a is supported; the results are illustrated in Figure 4. ANOVA analysis also revealed a significant effect of news fakery (vs. authenticity) on the sharing intention toward the news (F = 6.94, $p < 0.01$), and the mean value of sharing intention is larger in the real-news condition (mean = 3.506, S.D. = 0.089) than in fake-news condition (mean = 3.192, S.D. = 0.080). Thus, H1b is supported; the results are illustrated in Figure 22 and Figure 23.

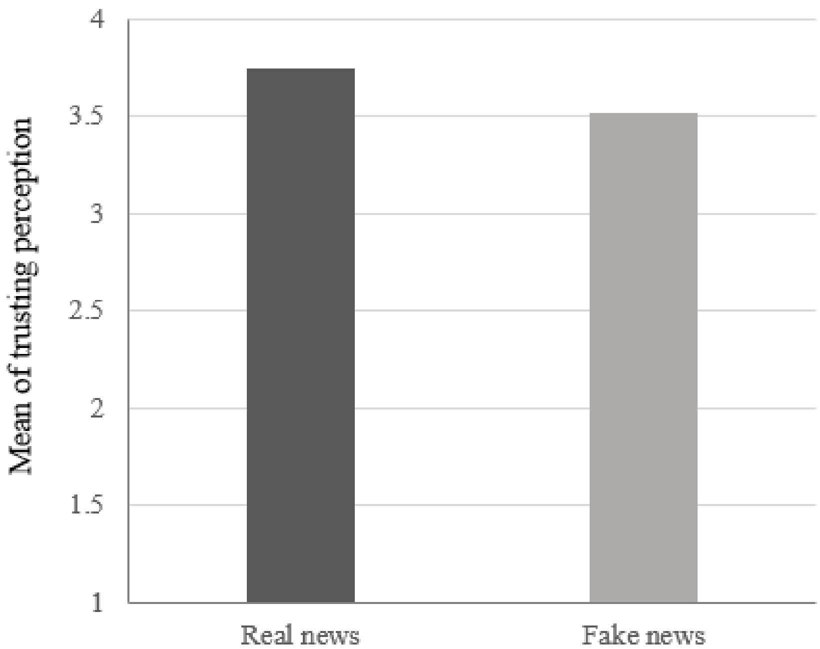


Figure 22 The effect of news fakery on trusting perception

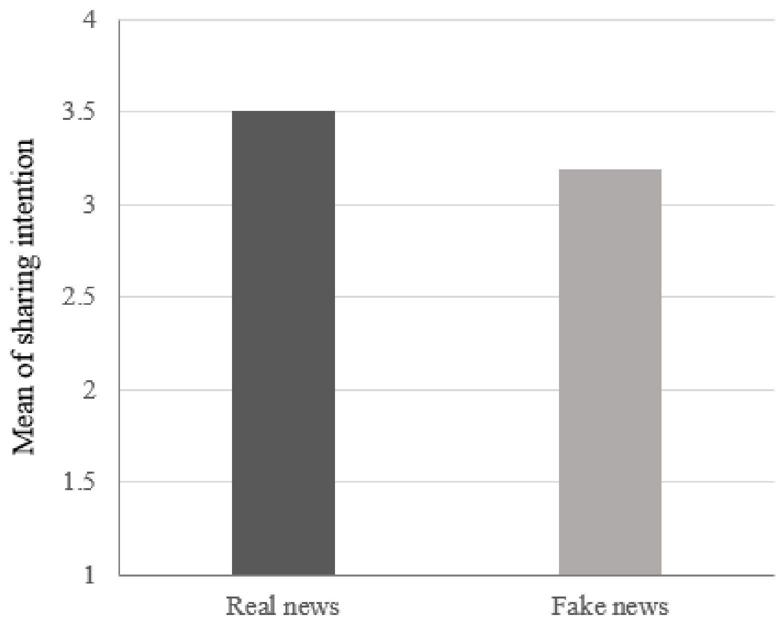


Figure 23 The effect of news fakery on sharing intention

We also tested H2, which posits that individuals' judgement accuracy toward real news and fake news decreases when news content triggers high arousal level compared to low arousal level. To be specific, individuals tend to show similar trusting and sharing intention toward fake news and real news with the existence of high-arousal content (H2a and H2b). We first employed ANOVA analysis to test whether the interactive item between news fakery and content arousal exerts a significant effect on trusting perception, which the results confirmed ($F = 6.90$, $p < 0.01$). Then, pairwise comparison was used to investigate the different effects of news fakery with high-arousal versus low-arousal content. The results indicated that subjects in the low-arousal-content condition had higher trusting perceptions toward real news (mean = 3.809) than toward fake news (mean = 3.319); the perception difference was significant ($t = -3.40$, $p < 0.01$). While subjects' trusting perception were similar toward both fake news (mean = 3.689) and real news (mean = 3.659) in the high-arousal-content condition ($t = 0.28$, $p > 0.1$). Thus, H2a is supported; the results are illustrated in Figure 24.

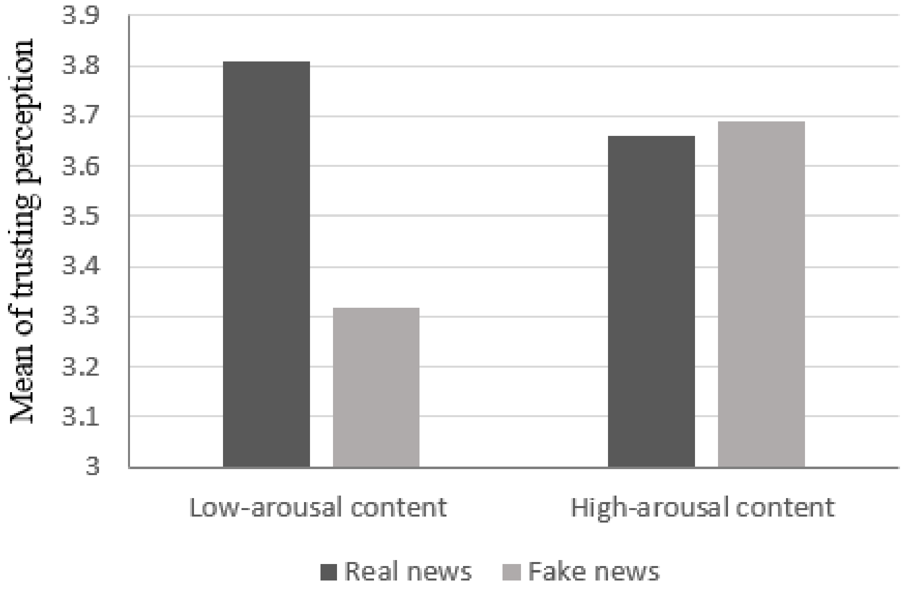


Figure 24 The moderating effect of arousal level of news content

We then employed ANOVA analysis to test the interactive effect between news fakery and content arousal on sharing intention, and the results confirmed a significant effect ($F = 2.92$, $p < 0.1$). Pairwise comparison was also used to investigate the different effects of news fakery with high-arousal versus low-arousal content. The results indicated that subjects in the low-arousal-content condition had higher sharing intention toward real news (mean = 3.511) than toward fake news (mean = 2.980); the intention difference was significant ($t = -3.10$, $p < 0.01$). Subjects' sharing intention were similar toward both fake news (mean = 3.333) and real news (mean = 3.511) in the high-arousal-content condition ($t = -0.72$, $p > 0.1$). Thus, H2b is supported (shown in Figure 25).

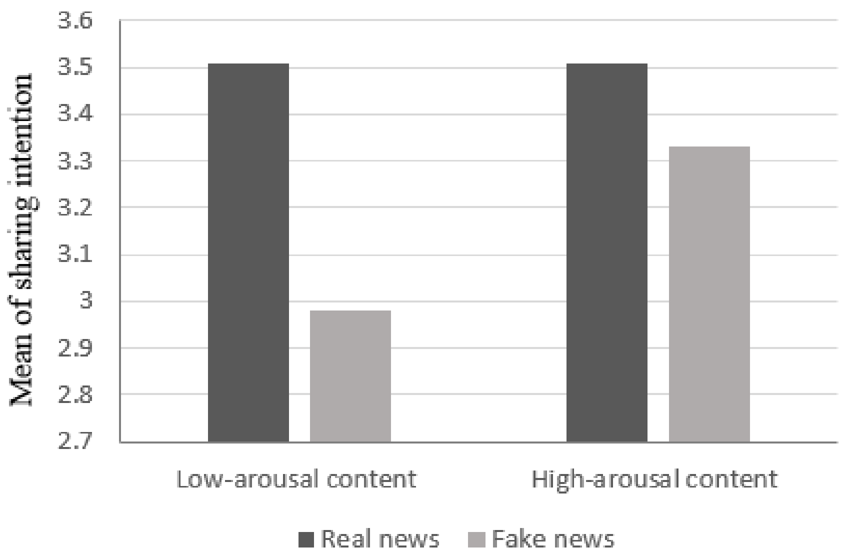


Figure 25 The moderating effect of arousal level of news content

H3 posits that individuals' personal involvement matters to their ability to distinguish between real news and fake news when the content triggers different arousal levels. That is, individuals who are highly involved in the epidemic event express similar trusting perception and sharing intention toward fake news and real news especially when the content triggers high arousal rather than low arousal (H3a and H3b). ANOVA analysis was conducted to test whether the interactive effect on trusting perception exists, which was confirmed by the results ($F = 3.25$, $p < 0.05$). Pairwise comparison was further used to investigate the different effects of news fakery with high-arousal versus low-arousal content when individuals have different personal involvement. The results indicated that for subjects who show less involvement toward the epidemic and belong to the low-arousal-content condition, they had higher trusting perceptions toward real news (mean = 3.500) than toward fake news (mean = 3.192); the perception difference was significant ($t = -1.79$, $p < 0.1$). Subjects' trusting perception were similar toward both fake news (mean = 3.333) and real news (mean = 3.250) in the high-arousal-content condition combined with low personal involvement ($t = 0.43$, $p > 0.1$). Furthermore, for subjects who are highly involved in the epidemic and belong to the low-arousal-content condition, trusting perceptions show significant differences ($t = -2.56$, $p < 0.05$) between real news (mean = 4.136) and fake news (mean = 3.600) with a larger value of trusting perception toward real news. However, for subjects who are highly involved in the epidemic and belong to the high-arousal-content condition, the difference between their trusting perceptions toward real news and fake news is reversed. That is, subjects' trusting perception toward fake news (mean = 4.222) is higher than that toward real news (mean = 3.885), and the perception difference was significant ($t = 1.88$, $p < 0.1$). Thus, H3a is supported; the results are illustrated in Figure 26.

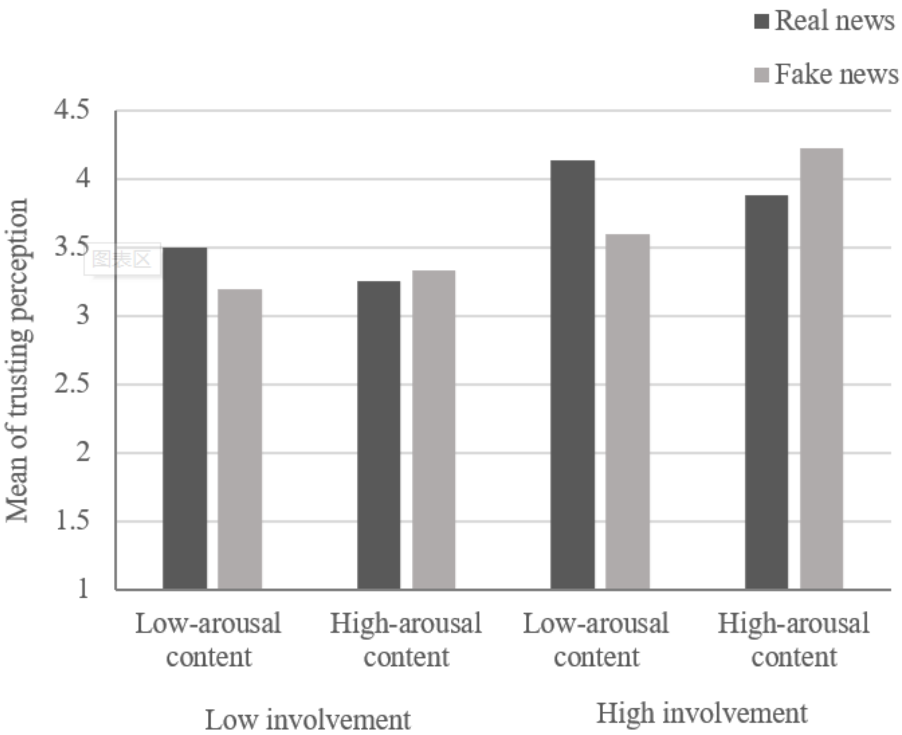


Figure 26 The moderating effect of personal involvement

We also conducted ANOVA analysis to test the existence of the interactive effect on sharing intention, and the results rejected the effect ($F = 1.41$, $p > 0.1$). Pairwise comparison was further employed and the results indicated that for subjects who show less involvement toward the epidemic and belong to the low-arousal-content condition, their sharing intention toward real news (mean = 3.000) than toward fake news (mean = 2.727) shows no significant difference ($t = -1.45$, $p > 0.1$). Subjects' sharing intention were also similar toward both fake news (mean = 3.000) and real news (mean = 3.000) in the high-arousal-content condition combined with low personal involvement ($t = -0.00$, $p > 0.1$). Furthermore, for subjects who are highly involved in the epidemic and belong to the low-arousal-content condition, sharing intention show significant differences ($t = -2.04$, $p < 0.05$) between real news (mean = 4.000) and fake news (mean = 3.533) with a larger value of trusting perception toward real news. However, for subjects who are highly involved in the epidemic and belong to the high-arousal-content condition, the difference between their trusting perceptions toward real news and fake news disappears. That is, subjects' sharing intention toward fake news (mean = 3.952) and toward real news (mean = 3.793) shows no significant difference ($t = 0.71$, $p > 0.1$). Thus, H3b is not supported; the results are illustrated in Figure 27.

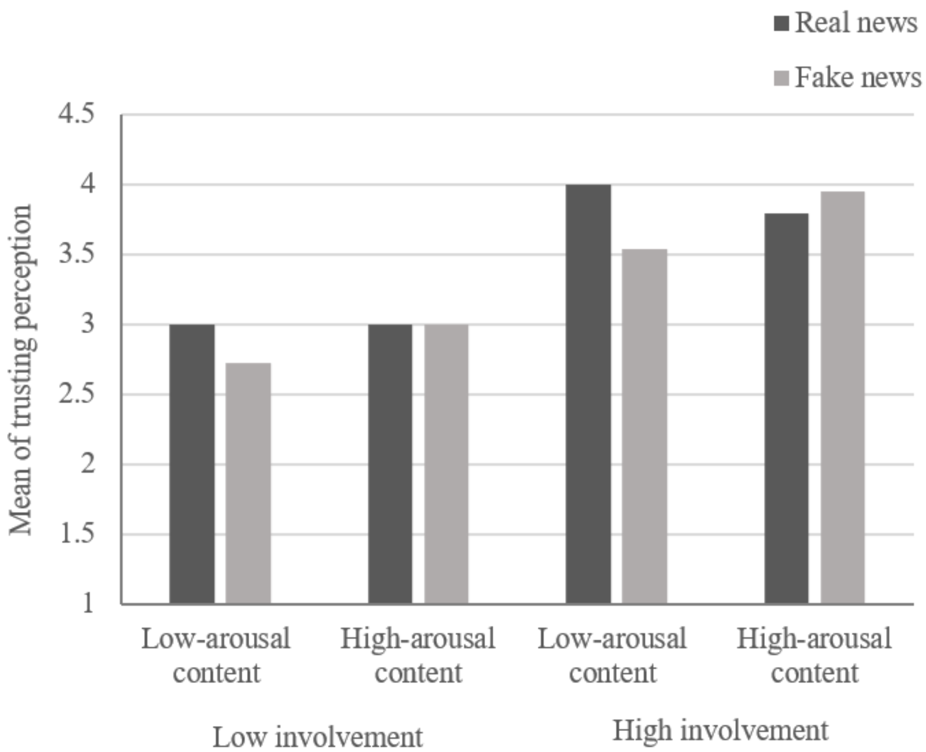


Figure 27 The moderating effect of personal involvement

3.6 Discussion

3.6.1 Findings

The results suggest that individuals have some basic ability to distinguish real news and fake news and arousal level and personal involvement are influential factors of individuals' intentions to trust and share toward real news and fake news. Specifically, individuals show higher trusting and sharing intention toward real news than fake news after ruling out other factors. This result is consistent with the findings in prior studies which identified individuals' intrinsic ability to discern fake news from real news in financial market (Clarke et al., 2019). One possible explanation is the linguistic differences such as language repetition and the usage of punctuation may exist between fake news and real news, making the content in real news less discountable (Busselle et al., 2000; Horne & Adali, 2017). As individuals keep rationality when making evaluations toward uncertain information (Doyle, 1992), they tend to think carefully toward news content and make rational judgement toward fake news and real news.

Two interesting findings also emerge. First, for individuals who are triggered a high arousal level by news content, their ability to discern fake news and real news becomes lower that their trusting perception and sharing intention toward fake news and real news show little difference. This finding is a consistent extension for existing studies investigating the positive effect of arousal level on information dissemination. In these existing studies, normal content (without considering content authenticity) who evoke high arousal level could be more viral as high arousal level can increase action-related behavior and reduce individual rationality (Gaertner & Dovidio, 1977; Berger & Milkman, 2012). Similarity, this study suggests that high arousal level increases the possibility that individuals form a narrow range of attention and an emotion-driven thinking. In this way, they may be less likely to process complex information and tend to ignore important aspects when making judgement toward news (Humphreys & Revelle, 1984; Gellatly & Meyer, 1992; Gernsbacher et al., 1998; Sanbonmatsu & Kardes, 1988). While for individuals who are triggered a low arousal level by news content, they can still distinguish between fake news and real news. Thus, with the existence of high arousal level, individuals have low ability to comprehend the news cent and discern news fakery.

Second, when individuals are highly involved in the epidemic event and are stimulated with a high arousal level by news content, their ability to distinguish fake news and real news becomes further lower that they even treat fake news as more trustworthy than real news. In marketing research, consumer involvement toward a brand has been identified as determinants of consumer engagement for the brand's activity (Liu & Jo, 2020). Our context-specific features that individuals cannot get away of the epidemic make it more important to explore the role of personal involvement. Consistent with prior studies, high involvement toward the epidemic event may induce individuals' engagement to focus on related news of the event. In this way, individuals tend to form a negative sentiment as such news mostly contain negative information. With the existence of such sentiment, individuals' cognitive level and decision quality toward news evaluation become lower (Lerner & Keltner, 2000, 2001). Combined with a high arousal level, individuals' capacity to judge news becomes even lower that they show higher trusting toward fake news rather than real news. While individuals do not have higher sharing intention toward fake news than real news even when they are highly involved in the epidemic event and have a high arousal level triggered by news content. The reason may be that although high involvement toward epidemic event can make individuals pay more attention and reduce rational thinking toward the news content, they only keep this irrationality inside their own judgements rather than sharing to others as trusting does not always lead to sharing (Seifert et al., 2017).

3.6.2 Theoretical implications

This study has some theoretical implications. Generally, this study extends existing fake news literature by exploring individual perceptions on fake news from the perspective of news content and individual features. Prior studies have already investigated how individuals perceive fake news with the usage of warning signs such as a flag of fake news (Mena, 2019; Moravec et al., 2018), the presentation of news source ratings (Kim et al., 2019), and the existence of critical opinions from peers (Colliander, 2019). While these studies mainly focused on the effect of platform designs on reader perceptions or judgements toward fake news. It's still unclear how news content and individual features work on the trusting and sharing intentions toward fake news. The implications of online randomized experiment help us rule out other factors and focus on the effect of news content and reader features on individuals' ability to discern fake news.

Specifically, this study contributes to arousal literature by introducing the influence of arousal level into the context of fake news. The negative influence of arousal has been identified in several areas including marketing (Gorn et al., 2001), child education (Valiente et al., 2012) and organizations (Pazzaglia et al., 2012). The rapid development of social media encourages scholars to focus on how arousal level works on information dissemination on such platforms (Okdie et al., 2013; Rubenking, 2019). However, the effect of individuals' arousal level also matters to news perceptions with the possibility that news is fake. This study adopts news content to manipulate individuals' low-arousal and high-arousal and explores the effect of arousal level on individuals' trusting and sharing intention toward fake news and real news, extending existing arousal literature to a new context.

In addition, this study complements the literature on personal involvement by introducing its concept to the context of epidemic event and testing its influence on individuals' perceptions on fake news. Existing studies on personal involvement mainly concentrated in marketing context and a positive relationship between product/customer involvement and customer loyalty as well as purchase intention is identified (Wu & Hsiao, 2017; Harun & Prybutok, 2020; McClure & Seock, 2020). Personal involvement toward epidemic event should be explored as epidemic event influences public life and stimulates public to involve in its related information. However, little studies considered the effect of personal involvement from this perspective. Inheriting from situation involvement by Houston and Rothschild (1978), this study defines epidemic event involvement as individuals' perception of epidemic event based on their inherent needs, values and interests and finds its significant moderating effect on individuals' evaluations toward fake news, thus contributing to existing personal involvement literature by investigating its effect in more contexts.

3.6.3 Practical implications

The findings in this study identify the bias of individuals when evaluating news about a specific epidemic event as news content and individuals' personal involvement can trigger individuals' own feelings and lead them to irrationality. Compared to that with low-arousal content (i.e., sad content), fake news with high-arousal content (i.e., angry content) is more deceptive, that is, is more likely to be perceived as trustworthy and gain sharing intentions from individuals. This situation is strengthened when the individuals has a high involvement toward the epidemic event. Thus, we suggest individuals to judge news with more rationality and more thinking before trusting or sharing a piece of news even though the individuals feel activated after reading the news or feel highly involved toward the news or the epidemic event. Otherwise, the individuals may trust fake news and even disseminate it to their friends.

For social media platforms, after realizing the bias of individuals, they could send warning messages or set warning signs to indicate the possibilities that the news takes advantage of the individuals' feelings and convince the individuals to keep an eye on the news. The platforms could also increase the regulations on fake news especially those contain high-arousal content (i.e., angry content) as such news can pose higher impact on reader perceptions. In addition, the platforms could pay more attention to individuals who are more involved in the news event. Specifically, for the most influential epidemic event (COVID-19), individuals from Hubei may deserve more attention as they are highly influenced by the epidemic and tend to have higher personal involvement to the event news, which may reduce their ability to discern real news and fake news.

3.6.4 Limitation and future studies

This study has some limitations. First, this study includes angry and sad content in news to manipulate different arousal levels to individuals. Although angry and sad content have been identified as efficient emotions to trigger individuals' arousal level, content with other emotions such as anxious and disgust can also work. This study only focuses on two types of emotions to simplify the experiments and investigate the research question more efficiently. Future studies could consider more content emotions and get a robust result. Second, this study employs five questions adapted from prior studies to measure personal involvement toward the epidemic event. As epidemic event such as COVID-19 influences people all around the country (or even world) but to varying degrees, the location (i.e., province) may be a more objective indicator to personal involvement. For example, participants from Hubei Province, China may be more involved to the epidemic. Future studies could collect data from a broader range to investigate the research question. Third, this study collects data from universities in China. COVID-19 is a global health crisis that individuals in other nations are also influenced and may also react to news referring to COVID-19. We encourage future studies to recruit participants from different nations to explore the impact of cultures on individuals' judgments toward fake news under different conditions. Fourth, this study conducts experiments using an online crowdsourced platform rather than using offline labs. Although this kind of experiments can reach to large populations, it cannot track respondents' specific features such as eye movement. Future experiments will be conducted in offline context to gather more dimensional data (i.e., eye movement data) to find an overall mechanism for all hypotheses. Last, four conditions in this study (i.e., fake news with angry content and real news with sad content) are designed with different pictures and titles. Results from such manipulations could only investigate the effect of pictures and titles without isolating the influence of content in news title. I will conduct more experiments to control news pictures and focus on the effect of news content.

Chapter 4
Conclusion

This study is one of the first to explore the factors that could influence individuals' perception on online disinformation. In ecommerce context, review rating, review content, and reviewer rating distribution can all contribute to individuals' perception on review manipulation. The usage of a mixed-method investigation (controlled experiment and field study) suggests that the effects of these factors are consistent in realistic ecommerce platform that review rating, review content, and reviewer rating distribution are closely related to the identification of fake reviews. Regarding to the news for COVID-19, arousal level triggered by news content and individuals' personal involvement can both affect individuals' ability to discern fake news and real news. That is, high arousal level induced by angry content in news will reduce individuals' ability to discern fake news. The negative influence of angry content is further strengthened if individuals are highly involved in the epidemic.

These findings open the gate for exploring online disinformation from the perspective of individual perception, which paves possible paths for future research. For example, future studies could follow the study to investigate other influential factors on individuals' ability to discern fake news such as the relevancy between news photo and news title and the transmission pattern of the news. In addition, as suggested by prior studies, only review content is not enough to identify fake review and reviewer features should get attention (Zhang et al., 2016). Rating distribution is an objective of a reviewer's rating history and habit, which may provide more value for fake review identification in future studies.

References

Aerts, G., Smits, T., & Verlegh, P. W. (2017). How online consumer reviews are influenced by the language and valence of prior reviews: A construal level perspective. *Computers in Human Behavior*, 75, 855-864.

Aghakhani, N., Oh, O., & Gregg, D. (2017). Beyond the Review Sentiment: The Effect of Review Accuracy and Review Consistency on Review Usefulness. *Proceedings of the 38th International Conference on Information Systems (ICIS)*, AIS.

Ahmed, H. (2017). Detecting opinion spam and fake news using n-gram analysis and semantic similarity (Doctoral dissertation). University of Victoria.

Ajao, O., Bhowmik, D., & Zargari, S. (2018, July). Fake news identification on twitter with hybrid cnn and rnn models. In *Proceedings of the 9th International Conference on Social Media and Society* (pp. 226-230).

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.

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.

Angrilli, A., Cherubini, P., Pavese, A., & Manfredini, S. (1997). The influence of affective factors on time perception. *Perception & Psychophysics*, 59, 972–982.

Ayeh, J. K. (2015). Travellers' acceptance of consumer-generated media: An integrated model of technology acceptance and source credibility theories. *Computers in Human Behavior*, 48, 173-180.

Ayeh, J. K., Au, N., & Law, R. (2013). "Do we believe in TripAdvisor?" Examining credibility perceptions and online travelers' attitude toward using user-generated content. *Journal of Travel Research*, 52(4), 437-452.

Balmas, M. (2014). When Fake News Becomes Real. *Communication Research*, 41(3), 430-454.

Banerjee, S., Bhattacharyya, S., & Bose, I. (2017). Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. *Decision Support Systems*, 96, 17-26.

Bapna, R., Goes, P., Wei, K. K., & Zhang, Z. (2011). A finite mixture logit model to segment and predict electronic payments system adoption. *Information Systems Research*, 22(1), p.118-133.

Barbado, R., Araque, O., & Iglesias, C. A. (2019). A framework for fake review detection in online consumer electronics retailers. *Information Processing & Management*, 56(4), 1234-1244.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173.

Barthel, M., Mitchell, A., & Holcomb, J. (2016). Many Americans believe fake news is sowing confusion. *Pew Research Center*, 15, 1-15.

Bentzen, N. (2019). Online disinformation and the EU's response. *European Parliamentary Research Service*. Retrieved from: https://epthinktank.eu/2018/04/24/online-disinformation-and-the-eus-response/.

Berger, J. (2011). Arousal Increases Social Transmission of Information. *Psychological Science*, 22(7), 891-893.

Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of Marketing Research*, 49(2), 192-205.

Bian, X., & Haque, S. (2020). Counterfeit versus original patronage: Do emotional brand attachment, brand involvement, and past experience matter? *Journal of Brand Management*, 1-14.

Bond, S. D., Bettman, J. R., and Luce, M. F. 2008. "Consumer Judgment from a Dual-Systems Perspective: Recent Evidence and Emerging Issues," in *Review of Marketing Research* (Volume 5), N. K. Malhotra (ed.), Armonk, NY: M. E. Sharpe.

Boush, D. M., Friestad, M., & Rose, G. M. (1994). Adolescent skepticism toward TV advertising and knowledge of advertiser tactics. *Journal of Consumer Research*, 21(1), 165-175.

Britwum, K., & Yiannaka, A. (2019). Consumer willingness to pay for food safety interventions: The role of message framing and issue involvement. *Food Policy*, 86(December 2018), 101726.

Brooks, A. W., & Schweitzer, M. E. (2011). Can Nervous Nelly negotiate? How anxiety causes negotiators to make low first offers, exit early, and earn less profit. *Organizational Behavior and Human Decision Processes*, 115(1), 43-54.

Burtch, G., Hong, Y., Bapna, R., & Griskevicius, V. (2018). Stimulating online reviews by combining financial incentives and social norms. *Management Science*, 64(5), 2065-2082.

Busselle, R. W., & Greenberg, B. S. (2000). The nature of television realism judgments: A reevaluation of their conceptualization and measurement. *Mass Communication & Society*, 3(2-3), 249-268.

Butler, E. A., Lee, T. L., & Gross, J. J. (2007). Emotion regulation and culture: Are the social consequences of emotion suppression culture-specific?. *Emotion*, 7(1), 30.

Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. R. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology*, 79(4), 644.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.

Cheliotis, L. K. (2020). Neither dupes, nor pipers: violent crime, public sentiment and the political origins of mass incarceration in the United States. *Current Issues in Criminal Justice*, 32(1), 1-21.

Chen, C. F., & Tsai, M. H. (2008). Perceived value, satisfaction, and loyalty of TV travel product shopping: Involvement as a moderator. *Tourism Management*, 29(6), 1166–1171.

Chen, P.-Y., Dhanasobhon, S., and Smith, M. D. 2014. All Reviews Are Not Created Equal: The Disaggregate Impact of Reviews and Reviewers at Amazon.Com. Working paper (available at http://ssrn.com/abstract=918083).

Cheung, C. M., Lee, M. K., & Rabjohn, N. (2008). The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities. *Internet Research*, 18(3), 229-247.

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9-38.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3), 345-354.

Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2012). Detecting automation of twitter accounts: Are you a human, bot, or cyborg?. *IEEE Transactions on Dependable and Secure Computing*, 9(6), 811-824.

Chua, A. Y., & Banerjee, S. (2018). Intentions to trust and share online health rumors: An experiment with medical professionals. *Computers in Human Behavior*, 87, 1-9.

Chung, C., & Darke, M. (2006). The consumer as advocate: Self-relevance, culture, and word-of-mouth. *Marketing Letters*, 17(4), 269-279.

Colliander, J. (2019). “This is fake news”: Investigating the role of conformity to other users’ views when commenting on and spreading disinformation in social media. *Computers in Human Behavior*, 97, 202-215.

Cowan, K., & Ketron, S. (2019). A dual model of product involvement for effective virtual reality : The roles of imagination , co-creation , telepresence , and interactivity. *Journal of Business Research*, 100, 483–492.

Cruz, R. E., Leonhardt, J. M., & Pezzuti, T. (2017). Second Person Pronouns Enhance Consumer Involvement and Brand Attitude. *Journal of Interactive Marketing*, 39, 104–116.

Danescu-Niculescu-Mizil, C., Kossinets, G., Kleinberg, J., & Lee, L. (2009). How opinions are received by online communities: a case study on amazon.com helpfulness votes. *Proceedings of the 18th International Conference on World Wide Web*, ACM, 141-150.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10), 1407-1424.

Delone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, 19(4), 9-30.

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., et al. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, 112(8), 2389–239

Doyle, J. (1992). Rationality and its roles in reasoning. *Computational Intelligence*, 8(2), 376-409.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66,183-201.

Eaton, L. G., & Funder, D. C. (2001). Emotional experience in daily life: Valence, variability, and rate of change. *Emotion*, 1(4), 413.

Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler, & J. Forgas (Eds.). Affect, cognition, and social behaviour: New evidence and integrative attempts (pp. 25–43). Toronto: Hogrefe.

Entman, R. M. (1989). How the media affect what people think: an information processing approach. *Journal of Politics*, 51(2), 347-370.

Erkan, I., & Evans, C. (2016). The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in Human Behavior*, 61, 47-55.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3), 223-241.

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, 52, 498-506.

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

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, 68(6), 1261-1270.

Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management*, 51, 174-185.

Filieri, R., Raguseo, E., & Vitari, C. (2018). When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type. *Computers in Human Behavior*, 88, 134-142.

Filieri, R., Raguseo, E., & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, 77, 333-341.

Fletcher, R., Cornia, A., Graves, L., & Nielsen, R. K. (2018). Measuring the reach of "fake news" and online disinformation in Europe. Reuters institute factsheet.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, 19(3), 291-313.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.

Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of White racism, arousal, and helping behavior. *Journal of Personality and social Psychology*, 35(10), 691.

Gao, B., Hu, N., & Bose, I. (2017). Follow the herd or be myself? An analysis of consistency in behavior of reviewers and helpfulness of their reviews. *Decision Support Systems*, 95, 1-11.

Gelfert, A. (2018). Fake news: A definition. *Informal Logic*, 38(1), 84-117.

Gellatly, I. R., & Meyer, J. P. (1992). The effects of goal difficulty on physiological arousal, cognition, and task performance. *Journal of Applied Psychology*, 77(5), 694.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, 12(2), 314-325.

Gernsbacher, M. A., Hallada, B. M., & Robertson, R. R. (1998). How automatically do readers infer fictional characters' emotional states?. *Scientific Studies of Reading*, 2(3), 271-300.

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.

Gohary, A., Hamzelu, B., & Alizadeh, H. (2016). Please explain why it happened! How perceived justice and customer involvement affect post co-recovery evaluations: A study of Iranian online shoppers. *Journal of Retailing and Consumer Services*, 31, 127–142.

Goodman, P. S., & Darr, E. D. (1998). Computer-aided systems and communities: Mechanisms for organizational learning in distributed environments. *MIS Quarterly*, 417-440.

Gorn, G., Tuan Pham, M., & Yatming Sin, L. (2001). When arousal influences ad evaluation and valence does not (and vice versa). *Journal of Consumer Psychology*, 11(1), 43-55.

Guadagno, Rosanna E, Rempala, Daniel M, Murphy, Shannon, & Okdie, Bradley M. (2013). What makes a video go viral? An analysis of emotional contagion and Internet memes. *Computers in Human Behavior*, 29(6), 2312-2319.

Guo, L., Su, C., & Lee, H. (2019). Effects of Issue Involvement, News Attention, Perceived Knowledge, and Perceived Influence of Anti-Corruption News on Chinese Students' Political Participation. *Journalism and Mass Communication Quarterly*, 96(2), 452–472.

Hann, I. H., Hui, K. L., Lee, S. Y. T., & Png, I. P. (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of Management Information Systems*, 24(2), 13-42.

Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin*, 36(11), 1576-1588.

Harun, A., & Prybutok, V. R. (2020). An investigation into the link between consumer's product involvement and store loyalty : The roles of shopping value goals and information search as the mediating factors. *Journal of Retailing and Consumer Services*, 52, 101933.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408-420.

Heilman, K. M. (1997). The neurobiology of emotional experience. *Journal of Neuropsychiatry*, 9, 439–448.

Hong, I. B. (2015). Understanding the consumer's online merchant selection process: The roles of product involvement, perceived risk, and trust expectation. *International Journal of Information Management*, 35(3), 322–336.

Hong, Y., Huang, N., Burtch, G., & Li, C. (2016). Culture, conformity and emotional suppression in online reviews. *Journal of the Association for Information Systems*, forthcoming, 16-020.

Horne, B. D., & Adali, S. (2017, May). This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *Eleventh International AAAI Conference on Web and Social Media*.

Houston, M. J., & Rothschild, M. L. (1978). Conceptual and Methodological Perspectives on Involvement. *Research Frontiers in Marketing: Dialogues and Directions*, 184–187.

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, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, 9(3), 201-214.

Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior*, 48, 17-27.

Huang, C. Y., Chou, C. J., & Lin, P. C. (2010). Involvement theory in constructing bloggers' intention to purchase travel products. *Tourism Management*, 31(4), 513–526.

Huang, K. Y., Chengalur-Smith, I., & Pinsonneault, A. (2019). Sharing Is Caring: Social Support Provision and Companionship Activities in Healthcare Virtual Support Communities. *MIS Quarterly*, 43(2).

Huang, L., Tan, C. H., Ke, W., & Wei, K. K. (2018). Helpfulness of Online Review Content: The Moderating Effects of Temporal and Social Cues. *Journal of the Association for Information Systems*, 19(6), 3.

Huang, N., Hong, Y., & Burtch, G. (2017). Social Network Integration and User Content Generation: Evidence from Natural Experiments. *MIS Quarterly*, 41(4), 1035-1058.

Humphreys, M. S., & Revelle, W (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, 91,153-184.

Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to Online Disinformation: A Framework for Cross-National Comparative Research. *The International Journal of Press/Politics*, 25(3), 493–516.

Hussain, S., Ahmed, W., Jafar, R. M. S., Rabnawaz, A., & Jianzhou, Y. (2017). eWOM source credibility, perceived risk and food product customer's information adoption. *Computers in Human Behavior*, 66, 96-102.

Jang, S. M., & Kim, J. K. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computers in Human Behavior*, 80, 295-302.

Jani, D., & Han, H. (2014). Personality, satisfaction, image, ambience, and loyalty: Testing their relationships in the hotel industry. *International Journal of Hospitality Management*, 37, 11-20.

Jin, Z., Cao, J., Zhang, Y., Zhou, J., & Tian, Q. (2016). Novel visual and statistical image features for microblogs news verification. *IEEE Transactions on Multimedia*, 19(3), 598-608.

Kankanhalli, A., Tan, B. C., & Wei, K. K. (2005). Contributing knowledge to electronic knowledge repositories: an empirical investigation. *MIS quarterly*, 113-143.

Keller, K. L., & Staelin, R. (1987). Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research*, 14(2), 200-213.

Kelley, H. H. (1967). Attribution theory in social psychology. In Nebraska symposium on motivation. University of Nebraska Press.

Kim, A., & Dennis, A. (2019). Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media. *MIS Quarterly*, 43(3), 1025-1039.

Kim, A., Moravec, P. L., & Dennis, A. R. (2019). Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings. *Journal of Management Information Systems*, 36(3), 931-968.

King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing*, 28(3), 167-183.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94(2), 211.

Kron, A., Pilkiw, M., Banaei, J., Goldstein, A., & Anderson, A. K. (2015). Are valence and arousal separable in emotional experience? *Emotion*, 15(1), 35.

Kumar, N., Venugopal, D., Qiu, L., & Kumar, S. (2018). Detecting review manipulation on online platforms with hierarchical supervised learning. *Journal of Management Information Systems*, 35(1), 350-380.

Lanzetta, J. T., Cartwright-Smith, J., & Eleck, R. E. (1976). Effects of nonverbal dissimulation on emotional experience and autonomic arousal. *Journal of Personality and Social Psychology*, 33(3), 354.

Lazer, D. M. J., Baum, M., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., et al. (2018). The Science of Fake News. *Science*, 359(6380), 1094–1096.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science*, 359(6380), 1094-1096.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-spescific influences on judgement and choice. *Cognition and Emotion*, 14(4), 473–493.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81, 146–159.

Li, N., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education*, 48(2), 301-317.

Li, S., Liu, Z., & Li, Y. (2020). Temporal and spatial evolution of online public sentiment on emergencies. *Information Processing and Management*, 57(2), 102177.

Lim, Y. S., & Van Der Heide, B. (2014). Evaluating the wisdom of strangers: The perceived credibility of online consumer reviews on Yelp. *Journal of Computer-Mediated Communication*, 20(1), 67-82.

Liu, J., & Jo, W. M. (2020). Value co-creation behaviors and hotel loyalty program member satisfaction based on engagement and involvement: Moderating effect of company support. *Journal of Hospitality and Tourism Management*, 43(February), 23–31.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412-3427.

Ma, W. W., & Chan, A. (2014). Knowledge sharing and social media: Altruism, perceived online attachment motivation, and perceived online relationship commitment. *Computers in Human Behavior*, 39, 51-58.

Malik, M. S. I., & Hussain, A. (2018). An analysis of review content and reviewer variables that contribute to review helpfulness. *Information Processing & Management*, 54(1), 88-104.

Martin, L., Sintsova, V., & Pu, P. (2014). Are influential writers more objective?: An analysis of emotionality in review comments. *Proceedings of the 23rd International Conference on World Wide Web*, ACM, 799-804.

Mayzlin, D. (2006). Promotional chat on the Internet. *Marketing Science*, 25(2), 155-163.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8), 2421-2455.

McClure, C., & Seock, Y.-K. (2020). The role of involvement: Investigating the effect of brand's social media pages on consumer purchase intention. *Journal of Retailing and Consumer Services*, 53, 101975.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, 13(3), 334-359.

Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology. Cambridge, MA: MIT Press.

Meinert, J., Mirbabaie, M., Dungs, S., & Aker, A. (2018, July). Is it really fake?–Towards an understanding of fake news in social media communication. In International Conference on Social Computing and Social Media (pp. 484-497). Springer, Cham.

Mena, P. (2019). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook. *Policy & Internet*. Doi: 10.1002/poi3.214.

Micu, A., Micu, A. E., Geru, M, & Lixandroiu, R. C. (2017). Analyzing user sentiment in social media: implications for online marketing strategy. *Psychology and Marketing*, 34(12), 1094–1100.

Mitchell, A. A. (1979). Involvement- potentially important mediator of consumer behavior. *Advances in Consumer Research*, 06, 191–196.

Moore, S. G. (2015). Attitude predictability and helpfulness in online reviews: The role of explained actions and reactions. *Journal of Consumer Research*, 42(1), 30-44.

Moravec, P., Minas, R., & Dennis, A. R. (2018). Fake News on Social Media: People Believe What They Want to Believe When it Makes No Sense at All. *Kelley School of Business Research Paper*, (18-87).

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34(1), 185-200.

Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013) What Yelp Fake Review Filter Might Be Doing. *Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013)*, July 8-10, 2013, Boston, USA.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220.

Okdie, Bradley M, Rempala, Daniel M, & Garvey, Kilian J. (2013). Life in the Bubble: Examining the forwarding of political videos. *Computers in Human Behavior*, 29(6), 2425-2430.

Ouellette, J., & Wood, W. (1998). Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior. *Psychological Bulletin*, 124(1), 54-74.

Pan, Y., & Zhang, J. Q. (2011). Born unequal: a study of the helpfulness of user-generated product reviews. *Journal of Retailing*, 87(4), 598-612.

Pazzaglia, F., Flynn, S., & Sonpar, K. (2012). Performance implications of knowledge and competitive arousal in times of employee mobility:"The immutable law of the ex". *Human Resource Management*, 51(5), 687-707.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879-891.

Qazi, A., Syed, K. B. S., Raj, R. G., Cambria, E., Tahir, M., & Alghazzawi, D. (2016). A concept-level approach to the analysis of online review helpfulness. *Computers in Human Behavior*, 58, 75-81.

Qi, J., Zhang, Z., Jeon, S., & Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. *Information & Management*, 53(8), 951-963.

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.

Racherla, P., & Friske, W. (2012). Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories. *Electronic Commerce Research and Applications*, 11(6), 548-559.

Richardson, J. T. (1980). Concreteness, imagery, and semantic categorization. *Journal of Mental Imagery*, 4(1), 51-58.

Richins, M. L., Bloch, P. H., Mcquarrie, E. F., Richins, M. L., & Bloch, P. H. (1992). How Enduring and Situational Involvement Combine to Create Involvement Responses. *Journal of Consumer Psychology*, 1(2), 143–153.

Rokeach, M. (1973). *The nature of human values.* Free Press.

Rozenkrants, B., Wheeler, S. C., & Shiv, B. (2017). Self-expression cues in product rating distributions: When people prefer polarizing products. *Journal of Consumer Research*, 44(4), 759-777.

Rubenking, B. (2019). Emotion, attitudes, norms and sources: Exploring sharing intent of disgusting online videos. *Computers in Human Behavior*, 96, 63-71.

Ruchansky, N., Seo, S., & Liu, Y. (2017, November). Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 797-806).

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology*, 85(2), 291.

Sanbonmatsu, D. M., & Kardes, F. R. (1988). The effects of physiological arousal on information processing and persuasion. *Journal of Consumer research*, 15(3), 379-385.

Schellekens, G. A., Verlegh, P. W., & Smidts, A. (2010). Language abstraction in word of mouth. *Journal of Consumer Research*, 37(2), 207-223.

Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour*, 11(3), 234-243.

Seifert, H. A., Malik, R. E., Bhattacharya, M., Campbell, K. R., Okun, S., Pierce, C., et al. (2017). Enabling social listening for cardiac safety monitoring: Proceedings from a DIA-CSRC co-sponsored think tank. *American Heart Journal*, 194, 107–115.

Seltzer, E. K., Horst-Martz, E., Lu, M., & Merchant, R. M. (2017). Public sentiment and discourse about Zika virus on Instagram. *Public Health*, 150, 170–175.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, 80(2), 159-169.

Shen, W., Hu, Y. J., & Rees Ulmer, J. (2015). Competing for attention: an empirical study of online reviewers' strategic behavior. *MIS Quarterly*, 39(3), 683-696.

Shin, S., Chung, N., Xiang, Z., & Koo, C. (2019). Assessing the impact of textual content concreteness on helpfulness in online travel reviews. *Journal of Travel Research*, 58(4), 579-593.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decision Support Systems*, 108, 1-12.

Sinha, A., & May, J. (2004). Evaluating and Tuning Predictive Data Mining Models Using Receiver Operating Characteristic Curves. *Journal of Management Information Systems*, 21(3), 249-280.

Spielmann, N., & Richard, M. O. (2013). How captive is your audience? Defining overall advertising involvement. *Journal of Business Research*, 66(4), 499–505.

Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, 14(1), 47-65.

Swann Jr, W. B., Gómez, A., Huici, C., Morales, J., & Hixon, J. G. (2010). Identity fusion and self-sacrifice: arousal as a catalyst of pro-group fighting, dying, and helping behavior. *Journal of Personality and Social Psychology*, 99(5), 824.

Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153.

Valiente, C., Swanson, J., & Eisenberg, N. (2012). Linking students' emotions and academic achievement: When and why emotions matter. *Child Development Perspectives*, 6, 129–135.

Verma, N., Fleischmann, K. R., & Koltai, K. S. (2018). Demographic factors and trust in different news sources. *Proceedings of the Association for Information Science and Technology*, 55(1), 524-533.

Visentin, M., Pizzi, G., & Pichierri, M. (2019). Fake news, real problems for brands: The impact of content truthfulness and source credibility on consumers' behavioral intentions toward the advertised brands. *Journal of Interactive Marketing*, 45, 99-112.

Wang, F., Liu, X., & Fang, E. E. (2015). User reviews variance, critic reviews variance, and product sales: An exploration of customer breadth and depth effects. *Journal of Retailing*, 91(3), 372-389.

Wang, H. C., Pallister, J. G., & Foxall, G. R. (2006). Innovativeness and involvement as determinants of website loyalty: II. Determinants of consumer loyalty in B2C e-commerce. *Technovation*, 26(12), 1366–1373.

Wang, H., Meng, Y., Yin, P., & Hua, J. (2016). A Model-Driven Method for Quality Reviews Detection: An Ensemble Model of Feature Selection. *Proceedings of the 15th Wuhan International Conference on E-Business (WHICEB)*, 573-581.

Waszak, P. M., Kasprzycka-Waszak, W., & Kubanek, A. (2018). The spread of medical fake news in social media–the pilot quantitative study. *Health Policy and Technology*, 7(2), 115-118.

White, Gregory L, Fishbein, Sanford, & Rutsein, Jeffrey. (1981). Passionate love and the misattribution of arousal. *Journal of Personality and Social Psychology*, 41(1), 56-62.

Willemsen, L. M., Neijens, P. C., Bronner, F., & De Ridder, J. A. (2011). "Highly recommended!" The content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication*, 17(1), 19-38.

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, 83(6), 1281-1297.

Wu, I. L., & Hsiao, W. H. (2017). Involvement, content and interactivity drivers for consumer loyalty in mobile advertising: The mediating role of advertising value. *International Journal of Mobile Communications*, 15(6), 577–603.

Wu, Y., Ngai, E.W.T., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, Forthcoming (available at https://doi.org/10.1016/j.dss.2020.113280).

Yin, D., Bond, S., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, 38(2), 539-560.

Yin, D., Mitra, S., & Zhang, H. (2016). Research note—When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Information Systems Research*, 27(1), 131-144.

Yin, Dezhi, Bond, Samuel D, & Zhang, Han. (2017). Keep Your Cool or Let it Out: Nonlinear Effects of Expressed Arousal on Perceptions of Consumer Reviews. *Journal of Marketing Research*, 54(3), 447-463.

Yoon, S., & Zhang, D. (2018). Social media, information presentation, consumer involvement, and cross-border adoption of pop culture products. *Electronic Commerce Research and Applications*, 27, 129–138.

Yoon, Y., Kim, A. J., Kim, J., & Choi, J. (2019). The effects of eWOM characteristics on consumer ratings: Evidence from TripAdvisor.com. *International Journal of Advertising*, 38(5), 684-703.

Zaichkowsky, J. L. (1986). Conceptualizing involvement. *Journal of Advertising*, 15(2), 4–34.

Zajonc, R. B. (1965). Social facilitation. *Science*, 149(3681), 269-274.

Zhang, D., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2), 456-481.

Zhang, J., Dong, B., & Philip, S. Y. (2020, April). Fakedetector: Effective fake news detection with deep diffusive neural network. In *2020 IEEE 36th International Conference on Data Engineering* (ICDE) (pp. 1826-1829). IEEE.

Zhang, W, & Watts, S. A. (2008). Capitalizing on content: Information adoption in two online communities. *Journal of the Association for Information Systems*, 9(2), 73-94.

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2), 102025.

Appendix. The Scale of Perceived Review Manipulation

Questions in Our Study	Original Item	Source
The review is disguised by related parties (e.g., the hotel itself or the competitor)	Firms can disguise their promotion as consumer recommendations due to the anonymity afforded by online communities.	Mayzlin (2006)
The review is not a truthful account of a real customer’s experience	Manipulation means that the posted review is not a truthful account of a real customer's experience.	Hu et al. (2012)
The review intends to mislead customers in their booking decision-making	Fake reviews were defined as deceptive reviews provided with an intention to mislead consumers in their purchase decision making.	Zhang et al. (2016)
The review is manipulated by related parties (e.g., the hotel itself or the competitor)	N/A	Self-developed