THE EFFECT OF CONSUMER ATTITUDINAL DISPOSITION IN ONLINE REVIEW KNOWLEDGE TRANSFER

THE EFFECT OF CONSUMER ATTITUDINAL DISPOSITION IN ONLINE REVIEW KNOWLEDGE TRANSFER

By MEHMET AKGÜL, B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements for the

Degree Doctor of Philosophy in Business Administration

McMaster University DOCTOR OF PHILOSOPHY IN BUSINESS ADMINISTRATION - INFORMATION SYSTEMS (2023) Hamilton, Ontario, Canada

TITLE: The Effect of Consumer Attitudinal Disposition in Online Review Knowledge Transfer

AUTHOR: Mehmet Akgül, B.Sc., M.Sc.

SUPERVISOR: Dr. Ali Reza Montazemi

NUMBER OF PAGES: xii, 141

Abstract

Online reviews, which are consumer-generated messages, play a vital role in the consumer decision making process especially prior to their purchase adoption (i.e., pre-usage). The objective of this research is to investigate the effects of two-sided online reviews' contents affecting the consumers' attitudes at the pre-usage stage of a focal experience service. Contrary to one-sided reviews (i.e., only positive or negative information), two-sided reviews contain both positive and negative information about a product/service: Two-sided reviews are considered more informative. Extant studies make an important assumption that there is no information asymmetry between writer/source of two-sided reviews and consumers that read/receive it. Their implicit assumption is that the attitude of the writer/source of the two-sided review is completely transferred to the reader/receiver of the review. Given the subjective nature of two-sided online reviews for experience goods, we contend that such an assumption is flawed because transfer of personal experience in form of attitude towards a focal object/service to others is fraught with ambiguity and uncertainty that can mitigate the transfer. Drawing on ambivalence and prospect theories, our hypothesis states that: the anticipatory ambivalence of the receiver/reader based on a two-sided review content for a focal service is higher than the ambivalent attitude of the source/writer of the review who has already experienced the focal service. Our empirical study, consisting of 1492 subjects from Canada and the United States, supports our stated hypothesis. The implication of our finding is profound. It shows that the extant literature had underestimated the negative attitude of the receiver/reader of the online reviews in their investigation, which confound their findings. To that end, we provide future research direction and implications of our findings in practice.

Acknowledgement

First of all, I wish to express my sincere appreciation to my Ph.D. supervisor, Dr. Ali Reza Montazemi, for his continuous support, help and guidance during my Ph.D. studies. Under his supervision, I not only excelled in scholarly work but also improved myself personally. This work is completed with his continuous support, which I cannot express my gratitude enough. I am also grateful to my Ph.D. supervisory committee members Dr. Maryam Ghasemaghaei and Dr. Manaf Zargoush, for their valuable support and feedback during my Ph.D. studies.

I would like to also thank my professors and friends in the Information System departments who helped me in this journey and DSB administration in guiding the Ph.D. process.

Finally, I would like to thank my dear parents Hacı Ali Akgül and Gülizar Akgül, my sister Demet Akgül and my brother Hasan Ali Akgül for their continuous support and help during my Ph.D. journey. It wouldn't be possible to complete this work without their support. I dedicate my dissertation work to them.

Table of Contents

List of Figures

List of Tables

1. Introduction

Consumers generally seek information as they are unsure about their decision to purchase a product/service (Berger, 2014). Although they can obtain information from formal information sources (e.g., brochures, handouts, newspapers, or advertising campaigns), they mostly rely on informal (i.e., personal information) sources before making a purchase decision (Bansal & Voyer, 2000). This informal communication is referred to as word of mouth (WoM) and is defined as "informal communications directed at other consumers about the ownership, usage, or characteristics of particular goods and services or their sellers" (Westbrook, 1987, p.261).

With the advances in information technology, especially web 2.0 technologies (e.g., consumer review sites or social networks), it became easier for people to share their opinion using information technology tools, and WoM is evolved into the online reviews (also called electronic word of mouth (eWoM)) and refers to as "any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet" (Hennig-Thurau et al., 2004, p.39). Case in point, Yelp, TripAdvisor, and Amazon where people post reviews to share their opinions about restaurants, hotels, or products, and the posts on the social networking services such as Facebook or Twitter.

Consumer review websites attract millions of consumers. For example, Yelp reached 265 million reviews for a variety of businesses (Yelp, 2023), and Facebook has around 2.96 billion monthly active users as of December 31, 2022 (Meta, 2023). Consumers rely increasingly more on informal information sources: "85% of consumers usually read up to ten reviews or opinions of other users before making their purchase decision" (Picher Vera et al., 2016, The Appearance of EWOM, para.3). Online reviews also affect product awareness; for example, by reading comments on Amazon.com, consumers may become aware of an unfilled product need and thus buy the product (Mudambi & Schuff, 2010). They also positively influence sales (Chen et al. 2008; Clemons et al. 2006), and 20% to 50% of purchasing decisions are derived by WoM (Bughin et al., 2010). In the United States, "31% of adults have rated a person, product, or service" (Zhang, Zhang et al., 2014) . In South Korea, 80% percent of the consumers refer to eWoM while searching for information about a product/service. (Doh & Hwang, 2009). Because of its impact on consumer decision-making, 64% of marketers recognize that eWoM messages are more effective than traditional marketing tools (Whitler, 2014).

1.1 Why people talk and what they talk about

Why do people share WoM, and why do some of them get talked/shared more than others? Berger (2014) provides answers to these questions using the five functions of WoM: Impression Management, Emotion Regulation, Information Acquisition, Social Bonding, and Persuading Others as follows.

Impression Management: One of the primary reasons people share word of mouth is to shape the impression they have on themselves, and others have on them since what they talk about impacts how others see them (Berger, 2014). Also, by sharing their experiences, people signal that they are knowledgeable in a specific area (Berger, 2014).

Emotion Regulation: WoM also facilitates consumers to regulate their emotions (Berger, 2014). Emotion regulation is "a person's active attempt to manage his emotional state by enhancing or decreasing specific feelings, or by reducing stress, anxiety or depression" (Shafir, 2015, p.231). People share their emotional experiences (Rimé, 2009), and generally, they share negative experiences to feel better (Berger, 2014).

In the case of negative experiences, people use WoM as a coping strategy because sharing their experience with others provides catharsis reducing emotional impact (Zech & Rimé, 2005). Consumers also use WoM as a venting mechanism when they are dissatisfied or angry (Wetzer et al., 2007). Although individuals are more inclined to share negative experiences, they also share positive ones as well. In such circumstances, sharing positive experiences enhances the positive effect of the experience (Gable et al., 2004).

Emotions convey not only valence (i.e., positivity or negativity) but also arousal as well. Bestelmeyer et al. (2017) define arousal (intensity) as "the level of autonomic activation that an event creates, and ranges from calm (or low) to excited (or high)" (p.1351). For example, although "anger, anxiety and sadness are all negative emotions" (Chou et al., 2022, p.4), sadness is a low arousal emotion (Raghunathan & Pham, 1999), and anger and anxiety are high arousal emotions (Yin, Bond et al., 2014). High arousal negative emotions increase the need to vent and hence they are shared more (Berger, 2014). On the positive side, due to its high willingness to experience it again (Berger, 2014), people share their positive experience through WoM.

Information Acquisition: As people are generally not sure about what they should do for a particular situation, they seek advice. In the consumer context, gathering information about products/services not only decreases the uncertainty about the purchasing decision but also associated perceived risk (Huang et al. 2007; Furner et al., 2013). Prospect theory shows that the effect of gains and losses on the individuals are different, and the psychological value of the loss is more than the gain (Kahneman & Tversky, 1979); hence consumers tend to minimize the loss (risk) through getting advice from others who already have purchased/used the pertinent products/services via WoM.

One way of information acquisition is to buy the product and try it (Lutz & Reilly, 1974). However, this is not only costly but also time-consuming. WoM plays a significant role in information acquisition as learning about other's experience (e.g., a negative customer experience from a specific brand) mitigates not only economic losses (e.g., trial and error by purchasing the product) but also time losses through obtaining information quickly and easily (Berger, 2014).

Social Bonding: Social relationships are a need (Thapa et al., 2022), and interpersonal communication such as WoM helps to fulfill that need (i.e., belonging to an online community) (Berger, 2014). Emotions also play a role in social bonding as high arousal emotions make individuals to connect with others (Berger, 2014).

Persuading Others: People use interpersonal communication to affect others in various domains, such as health behaviour and purchasing decisions (Berger, 2014). Although WoM provides the above functionalities, we would like to note that people may share their experiences and thoughts without realizing those functions. That is, they may share their experiences and thoughts as a result of intrinsic reward mechanism, as "sharing personal feelings and thoughts activates the same brain regions that respond to things like food, money, and seeing attractive members of the opposite sex" (Berger, 2014, p.597) and hence intrinsically rewarding (Tamir & Mitchell, 2012). This may explain why most social media posts, more than 70%, are about one's own personal experience (Berger, 2014).

1.2 WoM vs Online Reviews

Although WoM and Online Review provide the same functionality and serve the same purpose, there are slight differences between them, as outlined in Table 1. As can be seen from the table,

	WoM	Online Review
Credibility	The receiver of the information knows the communicator (positive influence on credibility)	Anonymity between the communicator and the receiver of the information (negative influence on credibility)
Privacy	The conversation is private, interpersonal (via dialogues), and conducted in real-time	The shared information is not private and, because it is written down, can sometimes be viewed by anyone and at any time
Diffusion speed	Messages spread slowly. Users must be present when the information is being shared	Messages are conveyed more quickly between users and, via the Internet can be conveyed at any time
Accessibility	Less accessible	Easily accessible

Table 1 Differences between WoM and Online Review (Adopted from Huete-Alcocer (2017))

while WoM is more personal, face-to-face communication, and hence more credible and private, online review is more accessible and can diffuse faster than WoM. As a result of its fast diffusion, Online Review is found to be more influential than WoM as Sun et al. (2006) state that "compared to traditional WoM, eWoM [Online Review] is more influential due to its speed, convenience, one-to-many reach, and its absence of face-to-face human pressure" (p.1106). Since online review is more accessible than WoM, Schiffman and Kanuk (2000) conclude that: "The expectation of receiving information that may decrease decision time and effort and/or contribute to the achievement of a more satisfying decision outcome" (p. 398).

Typologically, online reviews have two dimensions: communication scope and interaction level (Litvin et al., 2008). While the communication scope refers to the number of people included in an online review communication, e.g., one-to-one, many-to-one and many-to-many, the interaction level refers to whether the interaction between the participants is synchronous (simultaneous) or asynchronous (Litvin et al., 2008). Examples of online review communication for each category is depicted in Table 2.As can be seen from Table 2, online review websites are under the one-to-many, where one reviewer can share his/her thoughts/experiences with many other consumers simultaneously in an asynchronous way.

communication scope	Many-to-Many	Blogs and Virtual Communities	Newsgroups (e.g., Google Group)
	One-to-Many	Websites, Product Review and Hate Sites	Chatrooms
	One-to-One	Emails	Instant Messaging
		Asynchronous	Synchronous
		Level of Interactivity	

Table 2 Categorization of online review messages (Adopted from Litvin et al. (2008))

1.3 Information Load

Consumers have access to a vast number of reviews available on service provider sites, which brings additional challenges and costs. It is not only time-consuming to find and read the reviews, but the amount of information also makes it difficult for the consumer to process and judge reviews as a result of information load. Information load refers to "a complex mixture of the quantity, ambiguity and variety of information that people are forced to process. As load increases, people take increasingly strong steps to manage it" (Weick, 1995, p. 87). Roetzel (2019) modelled this information load and decision-making process as an inverted U curve, depicted in Figure 1. As shown in Figure 1, the decision-making performance ("the probability of achieving the best possible decision" (Roetzel, 2019, p.484)) increases as the information increases up to some point. After the optimal point, the performance decreases. This occurs for two reasons: cognitive capacity constraints and resource capacity constraints.

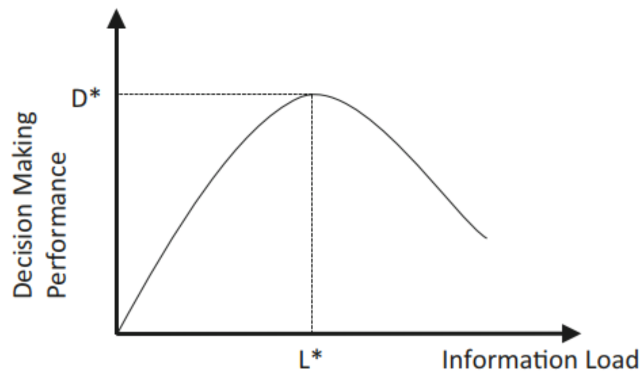

Figure 1 Information and decision-making performance (Roetzel,2019)

(a) Cognitive capacity constraints: This is an internal constraint to an individual; individuals have limited information processing capability (Simon, 1957); hence they cannot process information more than they can. This is also known as bounded rationality.

(b) Resource capacity constraints: This is an external constraint to an individual. This can happen when there are external constraints such as time and budget, and in such cases, decision-making performance decreases (Roetzel, 2019).

As consumers tend to decrease their effort while making a decision (Hu et al., 2014), they use various attributes of an online review, such as the numerical rating of the review to narrow down the choices (Racherla et al., 2012).

1.4 Online Reviews & Helpfulness

Review helpfulness refers to "the degree to which consumers perceive a product review to be helpful in their own purchasing decision-making" (Lopez & Garza, 2022, p.441). Thus, an online review's perceived helpfulness can be interpreted as measuring the online review's perceived value for the consumer (Mudambi & Schuff, 2010). There are a number of factors that affect review helpfulness as follows.

1.5 Factors Affecting Online Review Helpfulness

There has been extensive research in identifying factors that make a review helpful. Extant research shows that consumers rely on various attributes, such as star rating and the sender's reputation, to evaluate the online review (Qahri-Saremi & Montazemi, 2019). Since an online review is a part of the communication process between a sender of review and receiver of the review, the factors affecting the helpfulness of a review can be grouped into four as described by the Sender-Message-Channel-Receiver (SMCR) model (Berlo, 1960). The model assumes that there are four components in communication: Sender (S), Message (M), transmission medium (i.e., Channel (C)) and Receiver (R). Using the SMCR model for the content of the online review message, it can be inferred that:

- The source is the reviewer who has already purchased the products/services
- The message is the review text and associated review star rating
- The channel is the platform the review has been posted (e.g., blogs, social network sites or third-party consumer review websites such as Yelp.com or TripAdvisor.com), and
- The receiver of the message is the review reader (i.e., potential consumer)

Table 3 shows the factors affecting online review helpfulness from the extant literature. In this study, we controlled source, channel, and message-related attributes. Factors identified in Table 3 are used in extant literature affecting review helpfulness. Usually, this is done through a set of hypotheses. Demographic factors such as age, education, level of familiarity with online activities, and with online reviews used as control variables (i.e., no hypothesis tested) hence they are not included in Table 3.

Source	Channel	Message	Receiver
Source Credibility	Platform Type	Message Rating	Expertise
Source Expertise		Message Consistency	
Source Trustworthiness		Message Sidedness	
Source Attractiveness		Message Valance	
Gender		Message Quality	
		Message Credibility	
		Message Length	
		Message Readability	
		Message Style	

Table 3 Factors affecting online review helpfulness

1.5.1 Source Related Attributes

1.5.1.1 Source Credibility

Source credibility is defined as "the extent to which the source is perceived as possessing expertise relevant to the communication topic and can be trusted to give an objective opinion on the subject" (Goldsmith et al., 2000, p.44). Source credibility is not related to the message itself but is directly associated with the message owner. Studies show that information source credibility affects consumer information processing (Cheung et al., 2012). Source credibility is a complex concept, and according to the Source Credibility Model (Ohanian, 1990), it is identified by three components: source expertise, source trustworthiness and source attractiveness.

1.5.1.2 Source Expertise

Source Expertise refers to "the extent to which a communicator is expert, knowledgeable, experienced, qualified, or skilled" (Zhu et al., 2014, pp.270). Source expertise has a positive effect on the review helpfulness since the receiver of the review presumes that expert users are competent and provide correct information (Ismagilova et al., 2020; Zhu et al., 2014). Some platforms provide various attributes in an online environment to demonstrate the source's expertise level, such as the "Elite" badge on Yelp.com or the "Top 10,000 Reviewer" badge on Amazon.com.

1.5.1.3 Source Trustworthiness

Source trustworthiness refers to "consumers' perceptions that a source of communication is reliable, unbiased, and honest" (Filieri et al., 2018, pp.959). Studies show that source trustworthiness positively impacts the review helpfulness (Cao et al., 2018) as trustworthy sources are perceived to provide valid information. Although not directly observable, source trustworthiness can be inferred from the attributes provided in a review platform, such as the number of reviews submitted by a reviewer (Filieri, 2016) or the number of followers of a reviewer (Banerjee et al. 2017).

1.5.1.4 Source Attractiveness

Source attractiveness refers to the physical attractiveness of the source (Ohanian, 1990). Within an online community, the source attractiveness refers to the "online attractiveness" (i.e., the popularity of the source) in the community (Zhu et al., 2014). Zhu et al. (2014) used the number of the reviewer's online friends as a proxy for online attractiveness and found that online attractiveness and review helpfulness are positively associated.

1.5.1.5 Gender

The psychology literature shows that emotion norms are different between men and women, and generally, women express more emotions than men (Craciun & Moore, 2019). While women smile more and show more sadness, fear and guilts men express physically aggressive anger (Craciun & Moore, 2019). Moreover, depending on gender, people attribute different causes to emotional expressions. That is, while external factors (e.g., an event) are attributed to the men's emotional reactions, internal factors (e.g., personal characteristics) are attributed to the women's emotional reactions (Brescoll & Uhlmann, 2008). The attribution of emotional expressions affects the consumers' perception of the review (Kim & Gupta, 2012) and hence the review helpfulness. For example, Craciun et al. (2020) found that female-authored reviews are perceived as less helpful when expressing anger. Kwok and Xie (2016) also found that male reviewers' reviews are found to be more helpful compared to female reviewers' reviews.

1.5.2 Channel Related Attributes

1.5.2.1 Platform Type

Platform type refers to the platform where online review is shared. Consumers share their experiences and read others' experiences on various platforms such as retailers' websites and independent websites (Kwak et al., 2023; Lee & Koo, 2012). Consumers evaluate online reviews differently based on the platform they are posted (Kim et al., 2017; Lee & Youn, 2009). Due to the perceived credibility of the platforms, consumers rely more on the reviews that are posted on wellestablished and reputable websites or independent internet forums compared to the reviews posted on personal blogs or commercially oriented websites (Kim et al., 2017).

1.5.3 Message Related Attributes

1.5.3.1 Message (Review) Rating (Star Rating)

Message rating (star rating) shows the reviewer's overall evaluation of a product/service (Lopes et al. 2020), where 1 star represents an extremely negative experience, and 5 stars represent an extremely positive experience of the product/service (Mudambi & Schuff, 2010). Some of the extant studies show that star rating impacts (positively or negatively) the review helpfulness (Chatterjee, 2020; Chua & Banerjee, 2015)

1.5.3.2 Message (Review) Consistency

Message consistency indicates how the current review is consistent with the other reviews related to the same products/services (Cheung et al., 2009). It has been shown that there is a positive relationship between review consistency and message credibility. That is, the more a review is consistent (inconsistent) with the other reviews; it is perceived more (less) credible (Chakraborty & Bhat, 2018; Cheung et al., 2009) and hence found to be more helpful (Qahri-Saremi & Montazemi 2019).

1.5.3.3 Message (Review) Quality

Review quality "refers to the argument quality in a review message" (Shin et al., 2107, p.219). Argument quality is defined as "the audience's subjective perception of the arguments in the persuasive message as strong and cogent on the one hand versus weak and specious on the other" (Petty & Cacioppo, 1981, p. 264-265). Information with high-quality arguments

"(arguments that are relevant, objective and verifiable) tend to be viewed as more credible and persuasive" (Furner et al., 2013, p.425).

Review quality is positively associated with review helpfulness: as the quality of the argument increases, the receiver finds those reviews more logical and reliable (Chakraborty & Bhat, 2018), and those messages are found to be more helpful (Qahri-Saremi & Montazemi, 2019; Srivastava & Kalro, 2019).

1.5.3.4 Message (Review) Language Style

Although message quality has a positive effect on the review helpfulness, the use of language (i.e., writing style) can mitigate its effectiveness. Deceptive reviews contain more positive and less negative words (Yoo & Gretzel, 2009), and they tend to be verbose and ambiguous and longer compared to the authentic reviews (Banerjee & Chua, 2014).

1.5.3.5 Message (Review) Credibility

Message credibility refers to "the extent to which one perceives a recommendation / review as believable, true, or factual" (Cheung et al., 2009, p.12). Previous studies show that message credibility has a positive effect on the message helpfulness (Clare et al., 2018).

1.5.3.6 Message (Review) Length

Review length represents the extent of the information contained in a review. Previous studies show that longer reviews are more helpful as they contain more information about the product/service and its usage (Mudambi & Schuff, 2010; Filieri, 2016). They are also perceived to be more trustworthy than the shorter ones (Filieri, 2016).

1.5.3.7 Message (Review) Readability

Message readability refers to the understandability of the message. A message is most likely to be read if it is highly readable (Krishnamoorthy, 2015). Highly readable messages are also found to be more helpful (Aakash & Aggarwal, 2020; Korfiatis et al., 2012).

1.5.3.8 Message (Review) Valance

Review valence refers to the review's nature as positive, negative, or neutral (Kim et al., 2018). Prior studies show that consumers perceive positive and negative reviews differently and that negative reviews are the most influential ones (Zhang, Lee et al., 2010) and more helpful (Park & Nicolau, 2015). However, there are also other studies arguing that positive reviews, compared to negative reviews, are perceived to be more credible, trustworthy, and helpful (Pentina et al., 2018). The effects of message valence on the review helpfulness discussed through negativity bias, which states that individuals give more importance to negative information (Baumeister et al., 2001) as "negative information tends to be more diagnostic or informative than positive or neutral information" (Herr et al., 1991, p.460).

1.5.3.9 Message (Review) Sidedness

Review sidedness refers to the content of the review: "a one-sided review contains either positive or negative product comments, whereas a two-sided review contains both positive and negative comments on a product" (Cheung et al., 2012, p.622). Previous research found mixed results regarding the effect of review sidedness on the helpfulness of the review. While some studies show that one-sided reviews are more helpful compared to two-sided ones (Lee & Choeh, 2018; Chen, 2016), some studies found two-sided reviews more credible (Cheung et al., 2012) and are more helpful (Li, Lee et al., 2020; Filieri et al. 2018).

1.5.4 Receiver Related Attributes

1.5.4.1 Expertise

Expertise refers to the knowledge/experience related to the product/service (i.e., naïve vs expert). Depending on the product knowledge, experts and novices process information differently (Connors et al., 2011; Hong & Sternthal, 2010) as expert people perceive the two-sided reviews to be more helpful (Connors et al., 2011).

1.6 Research Objective

As stated before, online platforms, such as Yelp and TripAdvisor, provide convenient access to a large number of reviews (Qahri-Saremi & Montazemi, 2023). However, they impose an information load on the consumers due to their large number of reviews. It is not only time-consuming to find and read the reviews, but the amount of information also makes it difficult for the consumer to process and judge reviews as a result of the information load (Roetzel, 2019).

To mitigate the information load, it is recommended that review websites provide more useful information to consumers (Wang et al., 2020) by providing more helpful reviews, as review helpfulness is a measure of perceived value for the consumer (Mudambi & Schuff, 2010).

In the previous section, we identified the factors affecting online review helpfulness. However, for some factors, e.g., review star rating, review valence and review sidedness, the findings are mixed. Specifically for the review star rating, it is argued that it doesn't capture the reviewers' attitude accurately. For example, the star rating of a review, which is believed to measure a reviewer's attitude, is unable to identify indifference and ambivalence: individuals who are indifferent or ambivalent are inclined to give a 3 star to a product/service (Klopfer & Madden, 1980). Also, within these studies, it is implicitly assumed that the attitude enacted by the reviewer via the review text is completely transferred to the review reader. However, such an assumption may not be correct.

To the best of our knowledge, none of the studies investigated whether the review writers' attitudes are completely transferred to review readers. To ameliorate this void, in this study, we investigated the following research question:

Are the review writers' attitudes enacted in the review text completely transferred to the review readers?

In responding to the research questions, we will provide the theoretical background in the next chapter – chapter two. Then, in chapter three, we provide the details of the methodology of our research design, and in chapter four, we provide the analysis results. Finally, we discuss our findings in chapter five.

2 Theory Development

2.1 Information Flow

In an online review message, the reviewer conveys a message (the review) to the consumers. As the communication between the sender (the reviewer) and the receiver (the consumer) happens, the receiver judges the message and then makes inferences from that message. According to Berlo (1960), communication is a linear process where the information (message) flows from the sender to the receiver in a transmission medium. The model depicted in Figure 2 shows four components in communication: Sender (S), Message (M), transmission medium (i.e., Channel (C)) and Receiver (R).

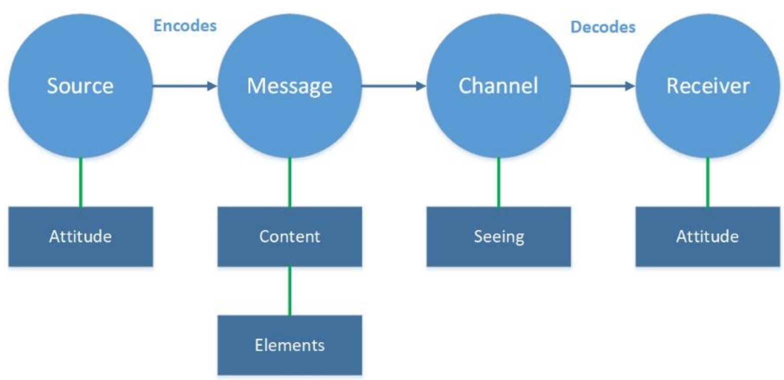

Figure 2 SMCR model for communication (Adopted from Berlo (1960))

The application of Berlo's (1960) model to online review communication is depicted in Figure 3. Tang and Guo (2015) contend that "[online] communication begins when an [online] sender develops attitudes [which can be positive, negative or two-sided] toward a product/service based on their consumption experience(s)" (p. 69). They then convert (encode) their thoughts and attitudes "into a text review of the product/service and an assigned star rating" (Tang & Guo, 2015, p.72) posted on consumer review websites (channel) such as TripAdvisor.com or Yelp.com, product review websites such as Amazon.com or social network sites. The receiver reads (seeing) the message and then creates an attitude towards the product/service.

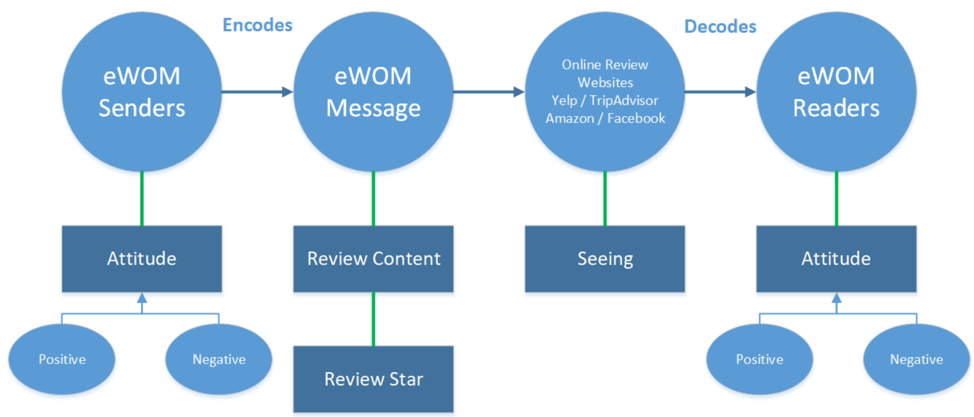

Figure 3 Information flow in online review communication (adapted from Berlo's (1960) SMCR model)

The objective of this research is to investigate the effects of two-sided online reviews' contents affecting the consumers' attitudes at the pre-usage stage of a focal experience service. Depending on the information content, online reviews can be classified into two categories: one-sided or two-sided. Contrary to one-sided reviews (i.e., only positive or negative information), two-sided reviews contain both positive and negative information about a product/service (Jensen et al., 2013; Uribe et al., 2016; Hsieh & Li, 2020). The receiver/reader of an online review shows their attitude towards the review by providing a helpful vote to the associated review. (For example, readers can press the "Helpful" button under the reviews on Amazon.com or use the thumbs-up emoji on TripAdvisor.com to show that they found the review helpful.) "Two-sided reviews are perceived to be more helpful" (Yin et al., 2021, p. 3) since they not only can be found to be more informative (Muralidharan et al., 2017), but also they reduce the receiver's skepticism (Cheung et al., 2009; Wang, Cunningham et al. 2015), enhance the reviewers' credibility (Jensen et al., 2013; Gerrath & Usrey, 2016), and increase the online review credibility (Cheung et al., 2012; Uribe et al., 2016; Park et al., 2019; Lopes et al. 2020).

Experience services cannot be objectively evaluated and can only be subjectively evaluated after consuming the service (Xiao & Benbasat, 2007). Examples of experience services include restaurant services and hotel services. Online reviews are one of the crucial information sources in reducing consumers' uncertainty in the evaluation of experience services (Qahri-Saremi & Montazemi, 2023). Therefore, we focus on the experience services for the context of this study. For example, restaurants offer "experiential service products which have the following attributes: intangibility (tangible products with intangible features such as taste and ambience), variability (quality inconsistency resulting from labor-intensive service delivery), and inseparability (simultaneity of production and consumption)" (Wang et al., 2021, p.1), thus making it difficult for customers to observe and evaluate quality of restaurant services before their first-hand experience creating uncertainty about service quality (Wang et al., 2021).

Uncertainty refers to "the degree to which the future states of the environment cannot be accurately anticipated or predicted" (Pfeffer & Salancik, 1978, p.67). In buyer-seller relationships uncertainty exists and perceived uncertainty is defined as "the degree to which the outcome of a transaction cannot be accurately predicted by the buyer due to seller and product related factors"(Pavlou et al., 2007, p.4). According to Pavlou et al. (2007) "uncertainty consists of seller quality uncertainty (seller hiding its true characteristics, making false promises, shirking, or defrauding), and product quality uncertainty (product condition not being as promised, or product quality being compromised)" (p.4). When consumers view service listing online, they may not have access to information about the "true" quality of the services and therefore, unable to judge service quality pre-usage (Pavlou et al., 2007).

The difference of information that sellers and buyers possess in the pre-usage stage refers to information asymmetry: sellers have more information than the buyers. Connelly et al. (2011) contend that "because some information is private, information asymmetries arise between those who hold that information and those who could potentially make better decisions if they had it" (p. 42). Information asymmetry "makes it difficult for buyers to assess the sellers' true characteristics and assess the true quality of a seller's products " (Pavlou et al., 2007, p.11) and "the higher the degree of information asymmetry that buyers perceive, the higher their uncertainty perceptions will be about a transaction" (Pavlou et al., 2007, p.11).

2.1.1 Online Review to Mitigate Information Asymmetry

To mitigate information asymmetry, consumers can use market information signals to distinguish seller's product/service quality (Pavlou et al., 2007). Online reviews, in terms of information signals, enable consumers to distinguish seller's product/service quality (Filieri, 2014). Mudambi and Schuff (2010) show that review helpfulness, "as a measure of perceived value in the decision-making process" (p.186), reflects information diagnosticity. The helpfulness of information plays a critical role in adopting the provided information as helpful information enables consumers to distinguish between alternative choices, making perceived helpfulness a fundamental predictor of information adoption (Qahri-Saremi & Montazemi, 2019). In the context of online review message adoption, extant research "focused on the perceived helpfulness of an online review message as an essential antecedent of online review adoption [since] … helpful online review messages reduce consumers' uncertainty in their assessment of a product/service."(Qahri-Saremi & Montazemi, 2019, p.8)

The foregoing research findings show that online reviews are instrumental in the form of information signals to mitigate information asymmetry between consumers at the pre-usage stage and the providers of services. Nonetheless, extant studies make an important assumption that there is no information asymmetry between writer/source of two-sided reviews and consumers that read/receive it: the attitude of the writer/source of the two-sided review is completely transferred to the reader/receiver of the review. Given the subjective nature of two-sided online reviews for experience goods, we contend that such an assumption is flawed because transfer of personal experience in form of attitude towards a focal object/service to others is fraught with ambiguity and uncertainty that can mitigate the transfer. Let us explain this phenomenon within the context of knowledge transfer in online communities (Faraj et al., 2016). To begin with, the personal experience (tacit knowledge) of the review provider is transferred to explicit knowledge in the form of review content. Next, the receiver of the reviews interprets and "incorporates them into their own personal tacit knowledge [i.e., internalize]" (Faraj et al., 2016, p. 675). Faraj et al. (2016) explained this process as:

> In online communities, such flows capture participants' evolving interpretations, understandings, and practicing of explicit knowledge, picked up and refined from the knowledge system [e.g., Yelp restaurant reviews] … [Information] flows continuously through the community and individuals [i.e., consumers at the pre-usage stage] need to select and incorporate them into their own personal tacit knowledge (p.675).

Expansive flow of information in online communities presents cognitive challenges for individuals to find relevant information, as they must conserve scarce human attention on evolving knowledge systems (Faraj et al., 2016). Extant literature identified allocation of attention processes in online behaviors. For example, Browne et al. (2007) found that for well-structured tasks "[tasks] that are of low or medium complexity and for which people have at least some experience" (p.98), individuals generally use mental list rule or the single criterion rule; and for poorly-structured tasks "[tasks] that are of high complexity and for which people have little or no experience [such as a fresh visit to a restaurant]" (Browne et al., 2007, p.98) individuals generally use magnitude threshold and representational stability rules. As Browne et al. (2007) contend, "the dimensions of the task structure and the nature of the person's representation [of the task]" (p.98) have a significant impact in the selection of the stopping rule during the information search. For example, for tasks such as selection of a search good, e.g., a notebook, the decomposition strategy is used since the information is discrete "various task elements, criteria, or attributes can be separately identified" (Browne et al., 2007, p.92). On the other hand, for holistic tasks (e.g., selecting an experience good) where information may not be discreet an individual acts "based on his 'sense' or 'image' or 'gist' of the situation rather than on individual elements" (Browne et al., 2007, p.93). In the latter case, there is an information asymmetry between the source and receiver of the reviews: the receiver of the review is limited to what the source disclosed in the review (Siddiqi et al., 2020).

As discussed before, information asymmetry increases a consumer's perceived risk when buying products/services (Pavlou et al., 2007). Perceived risk is "the consumer's perception of the uncertainty and concomitant adverse consequences of buying a product or service" (Dowling &

Staelin, 1994, p.119). Furthermore, the sender and receiver of the review have two different prospect-based emotions. The prospect-based emotions are "characterized as reactions to (i.e., having a positive or negative feeling about) an envisaged [anticipatory] event [by the reader of the review], or to the confirmation/disconfirmation of the prospect of such an event [by the writer/sender of the review]" (Ortony et al., 2022, p.126). While both confirmation/disconfirmation emotions and anticipatory emotions are based on assessments of future events, the difference between the two is that the former is the result of "a comparison between an actual outcome [of the event] and the expectation [about the event before the event happens]" (Bee & Madrigal, 2013, p.379); the latter is about the future events where the status of the event is unknown (Bee & Madrigal, 2013). To that end, Bee and Madrigal (2013) found that "anticipatory mixed [positive/negative] emotions (i.e., anticipatory ambivalence) … were the greatest contributor to consumer discomfort and uncertainty regarding future consumption [i.e., post-usage stage]" (Bee & Madrigal, 2013, p.370). They also assessed the effect of anticipatory ambivalence on attitudes and intentions for prospective purchase intentions and they found that "ambivalence mediates the relationship between evaluative information and intentions, as well as [ambivalence] moderates the relationship between attitudes and intentions" (p. 370).

2.2 Ambivalence

Ambivalence refers to "a state in which individual experiences both positive and negative reactions to an attitudinal object" (Yang & Unnava, 2016, p. 332). For example, a person has a positive attitude towards an object (e.g., French fries for taste), but at the same time they have some negative attitude towards the same object (e.g., high calorie of French fries), making them overall attitude ambivalent (Yang & Unnava, 2016). In the case of no conflicting evaluations (i.e., positive (e.g., liking the French fries without any reservation) or negative dominant (e.g., not liking French fries at all) evaluations), then the attitude is called univalent (Yang & Unnava, 2016; Chang, 2011). The difference between non-ambivalent (univalent) attitudes and ambivalent attitudes is that "non-ambivalent attitudes are generally based on evaluatively congruent attributes while ambivalent attitude-holders need to integrate evaluatively incongruent attributes into an overall judgment" (van Harreveld, 2004, p.431).

We can show the relationship between review content and ambivalence attitude in terms of heuristic-systematic model. The heuristic-systematic model highlights differences in informational processing (Qahri-Saremi & Montazemi, 2019), indicating that message recipients elaborate on messages differently in different situations (Chaiken, 1980). The heuristic-systematic model identifies two modes of information processing: systematic processing mode and heuristic processing mode (Chaiken, 1980). Systematic processing mode is the effortful processing mode, and it is used, for example, when the quality of the message arguments is carefully examined in the assessment of message helpfulness (Chaiken, 1980). Drawing on heuristic-systematic model, Qahri-Saremi and Montazemi (2019) show that "when a consumer is able and willing to engage in the systematic processing of a message, the merits of the actual arguments contained within the message serve as systematic cues to determine the degree of informational influence" (Qahri-Saremi & Montazemi, 2019, p.9). Extant research show that ambivalence is linked to enhanced systematic processing (van Harreveld et al., 2015) due to the "motivation to reduce ambivalence, as it was found that the increased receptiveness of ambivalent attitude holders to a strong persuasive message helped to reduce subsequent feelings of ambivalence" (van Harreveld et al., 2015, p.19). Ambivalence can be reduced using the unbiased systematic processing (van Harreveld et al., 2015) which is the "effortful processing achieved by carefully weighing all alternatives [in the review content] aiming to come to the best possible evaluation [of the focal service]" (van Harreveld et al., 2015, p.20).

2.2.1 Measuring Ambivalence

We can identify two central elements of ambivalence (van Harreveld et al., 2015): "First, both positive and negative associations need to be present. Second, these associations can be relevant at the same time. Based on these two prerequisites, we can make a distinction between the associative structure of ambivalence based on positive and negative association weights (objective ambivalence) and the experience of conflict due to this associative structure (subjective ambivalence)" (Van Harreveld et al., 2015, p.4).

2.2.1.1 Subjective Ambivalence

Subjective ambivalence, also known as felt ambivalence, can be measured using self-report questions. The subjective ambivalence measure developed by Priester and Petty (1996) consists of three items asking participants to rate (on an 11-point scale) how much their reactions "are conflicted, mixed, and indecisive to the attitude objects" (Priester & Petty, 1996, p.437). The subjective ambivalence is calculated by averaging one's responses to these three items.

However, people may not be able to accurately assess their subjective ambivalence (Has et al., 1992; Ullrich, 2012). For example, people could use ambivalence "to express their feelings of indecision or uncertainty" (Larsen et al., 2001, p.692) leading to inaccurate measurement of the subjective experience of being ambivalent (Russel et al., 2011): "people may hold ambivalent views about an issue but are unable to explicitly express those views as divided because the conflict is unconscious "(Russel et al., 2011, p.359). To that end, Russel et al. (2011) findings highlight the challenges in the measurement of subjective experience of being ambivalent and the importance of the use of indirect measures (i.e., measuring positive and negative evaluations separately) and the use of a computational formula in the measurement of ambivalence. Furthermore, McGraw et al. (2003) found that while "the information foundations of subjective ambivalence and uncertainty yielded in mixed results, … the objective measures of ambivalence did predict subjective uncertainty" (p. 435). Therefore, we use objective ambivalence to assess the consumers' attitude towards a focal service.

2.2.1.2 Objective Ambivalence

Objective ambivalence measures ambivalence using a mathematical formula. In this case, an individual is asked (e.g., "to what extent does each of the words below describe your attitude toward [topic]?" (Weng & DeMarree, 2019, p.4)), to rate their positive and negative attitude towards an entity (e.g., "a person, situation, object, task, or goal" (Rothman et al., 2017, p. 33)), and then those evaluations are used to compute the ambivalence (Breckler, 1994).

One of the formulas used to measure Objective Ambivalence is Griffin's ambivalence measure, also known as the SIM (Similarity-Intensity model) (Thompson et al., 1995). According to the model, ambivalence has two components: Intensity Component and Similarity Component

Intensity Component: The intensity component of ambivalence measures how strong the magnitude of the conflicting assessment is. The formula for intensity component is $(P+N)/2$, where P and N represent the positive and negative evaluations of an entity, respectively.

Similarity Component: The similarity component of ambivalence measures how similar in magnitude the positive and negative evaluation of an entity is. The formula for similarity is -

ABS(P-N), where P and N represent the positive and negative evaluations of an entity, respectively, and ABS refers to the absolute value function. The combined formula for ambivalence is provided in Equation 2:

$$Ambivalence = \frac{P + N}{2} - abs\ (P - N) \tag{2}$$

A 3D graph of the SIM model and associated contour plot is depicted in Figure 4. In the graph, it is assumed that the scale for the evaluations of Positive and Negative components is between 1 and 5. As Figure 4 shows if the similarity is kept constant, meaning that the positive and negative evaluations have the same strength, ambivalence increases as the evaluations' strengths increase. Also, suppose the intensity is kept constant, meaning that the sum of positive and negative evaluations is kept constant. In that case, ambivalence decreases as the distance between positive and negative evaluations increases, meaning that the evaluation becomes either positive or negative dominant.

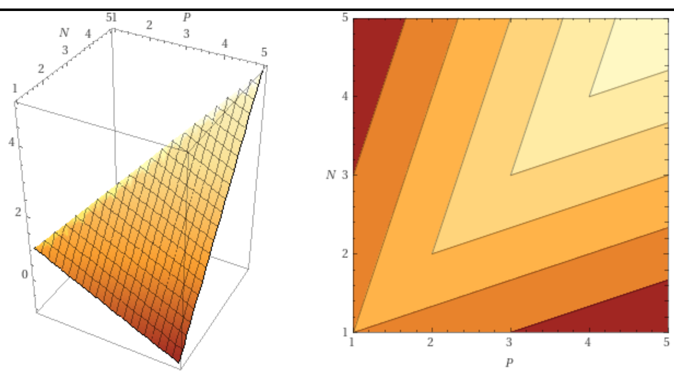

Figure 4 3D (left) and contour graphs (right) of ambivalence as modelled by SIM. As the color becomes lighter ambivalence increases.

As aforementioned, SIM contends that ambivalence has two components – intensity component and similarity component – where both the intensity and the similarity are calculated using the positive and negative evaluations of a stimulus (i.e., a person, an organization, or an experience). A heatmap of equation 2 is provided in Figure 5, where Positive and Negative evaluations are rated between [0, 9]. Each cell in Figure 5 shows the value of ambivalence for that specific evaluation: for example, the ambivalence of feeling 2 Positive (P=2) and 4 Negative (N=2) is 1 $(= \frac{2+4}{2} - abs(2-4) \; = 3 - 2 \; = 1)$.

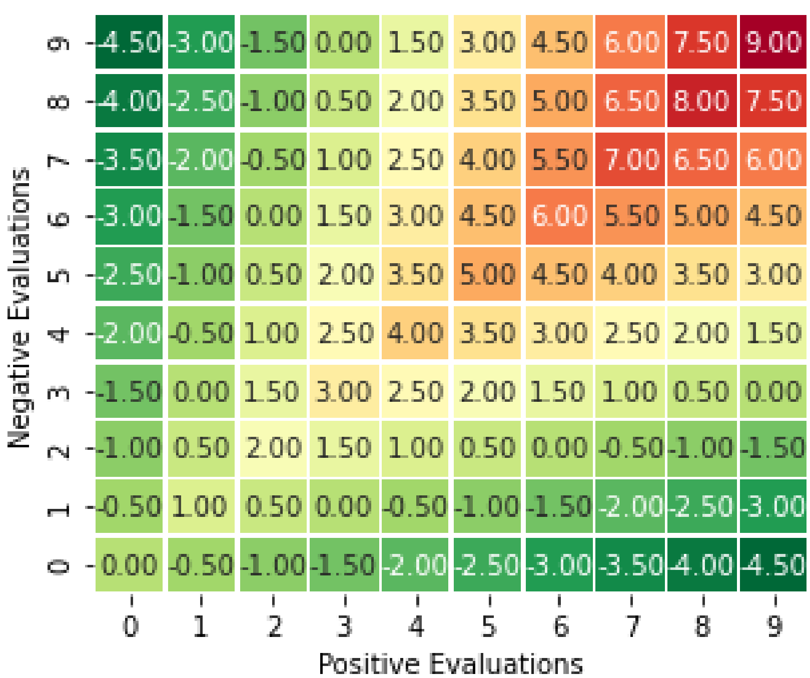

Figure 5 Heatmap of ambivalence calculated using Equation 2. While the x and y-axis show the positive and negative evaluations, the values in the cells show the ambivalence value for that specific evaluation. Ambivalence increases, going from green colour to yellow colour to red colour.

The first part of the SIM formula – (P+N)/2 – shows the **intensity component** of the ambivalence, and it represents that as the total intensity of the emotions increases, the ambivalence increases. This can be observed on the secondary diagonal (diagonal from bottom-left to top-right) in Figure 5. On the secondary diagonal Positive and Negative evaluations are the same, making $abs(P-N)$ to be 0. On the secondary diagonal, as the intensity of the evaluations increases, ambivalence increases and reaches its maximum value of 9 when both Positive and Negative evaluations become 9 (top-right corner in Figure 5).

The second part of the SIM formula – $abs(P-N)$ – is the **similarity component** and represents the distance between Positive and Negative evaluations. As the distance between Positive and Negative evaluations increases, ambivalence decreases, making the overall evaluation more univalent rather than ambivalent. For example, in Figure 5, ambivalence is minimum (=-4.5) when the Positive evaluation is 9, and the Negative evaluation is 0 (bottom-right corner in Figure 5) or when the Negative evaluation is 9 and the Positive evaluation is 0 (top-left corner in Figure 5). At these points, attitudes become univalent since there is no coexistence of mixed evaluations. The effect of the similarity component can be observed in Figure 5 by looking at the values on the main diagonal (diagonal from top-left to bottom-right). On the main diagonal, the sum of Positive and Negative evaluations is the same, making (P+N)/2 constant (=4.5). As shown on the main diagonal, ambivalence increases as the Positive and Negative evaluations become closer to each other and reaches its maximum value of 3.5 when the Positive evaluation is 5 and the Negative evaluation is 4 or when the Negative evaluation is 5, and the Positive evaluation is 4.

2.3 Hypothesis Development

As stated before, the sender and receiver of the review have two different prospect-based emotions. The prospect-based emotions are "characterized as reactions to (i.e., having a positive or negative feeling about) an envisaged [anticipatory] event [i.e., anticipatory ambivalence by the reader of the review], or to the confirmation/disconfirmation of the prospect of such an event [i.e., ambivalent attitude) by the writer/sender of the review]" (Ortony et al., 2022, p.126). While both confirmation/disconfirmation emotions and anticipatory emotions are based on assessments of future events, the difference between the two is that the former is the "result of a comparison between an actual outcome [of an event] and an expectation [of the event before it happens]" (Bee & Madrigal, 2013, p.379); the latter is about the future events where the status of event is unknown (Bee & Madrigal, 2013).

Drawing on prospect theory, we contend that after reading a review, the anticipatory ambivalence attitude of the receiver of the review is higher than the ambivalent attitude of the source/writer of the review who has already experienced the focal service. Prospect theory (Kahneman & Tversky, 1979) posits that individuals place more weight on the experience of loss than the pleasure of obtaining an amount equal to what was lost since the value function is steeper compared to the gains. That is, potential loss rather than prospective gain weighs more heavily on the decision between two possible options (Park & Nicolau, 2015). Furthermore, "according to the cognitive dissonance theory … people immediately focus their attention on unfavorable aspects of the chosen alternatives and favorable aspects of the rejected alternatives" after a decision (Van Harreveld et al., 2015, p. 11). Additionally, compared to inactions, actions are linked to higher degrees of regret and as making a decision requires action, it is more likely that the ambivalent attitude holder to feel regret after making a decision (Van Harreveld et al., 2015). As a result, when the holder of an ambivalent attitude is forced to make a decision, ambivalence is felt to be particularly unpleasant (van Harreveld et al., 2009).

Therefore, based on the foregoing justifications, we postulate the following hypothesis:

Hypothesis: The anticipatory ambivalence of the receiver/reader based on a two-sided review content for a focal service is higher than the ambivalent attitude of the source/writer of the review who has already experienced the focal service.

3 Methodology

To test the stated hypothesis, we conducted a controlled laboratory experimental study. For the controlled laboratory experiment, we used two sets of participants: Amazon Mechanical Turk (MTurk) participants and student participants enrolled in the "Information Systems in Business" course at the DeGroote School of Business. The rationale is to increase our findings' overall validity and generalizability; using different subjects will also reduce false interpretations and increase the strength of our findings (Hales, 2010). To that end, we used four scenarios, depicted in Table 4, to assess the transfer of knowledge (emotions/attitudes) from the review writers embedded in their reviews to the receiver. Scenario 1 consisted of 90 Canadian review readers and 90 Canadian/US review writers. Scenario 2 consisted of 157 US review readers and 135 Canadian/US review writers. Scenario 3 consisted of 321 student review readers and 228 Canadian/US review writers. Scenario 4 consisted of 314 student review readers and 240 student review writers. We provide the details of the controlled laboratory experiment in the following sections.

Scenarios	Knowledge Transfer from Review Writers to the Readers
Scenario 1	Canadian MTurk review reader – Canadian/US MTurk review writer
Scenario 2	US MTurk review reader – Canadian/US MTurk review writer
Scenario 3	Student review reader – Canadian/US MTurk review writer
Scenario 4	Student review reader –Student review writer

Table 4 Data collected to assess the stated hypothesis

3.1 Study Design, Context, and Subjects

In this study, two phases of self-administrated online surveys were conducted. In the first phase, online reviews about restaurants were collected from one subset of the participants, i.e., review writers. In the second phase, written online reviews are shown to another subset of participants, i.e., review-readers. We collected the data using online surveys as online surveys have been "particularly advantageous for social studies in terms of reduced costs, immediacy and enhanced questionnaire possibilities" (Ferri-García & Rueda, 2022, p.1829). To ensure that the survey questions were understandable to a broad audience, we also conducted a pilot study before starting the controlled experiment.

3.1.1 Study Context

For the study context, we selected experience services as defined in the Theory Development chapter. In general, services can be classified as "search" or "experience" services (Mitra et al., 1999). Search services can be objectively evaluated without the need for the consumers to experience them (Xiao & Benbasat, 2007). Examples of search services include checking account or credit card for banks (Licata & Frankwick, 1996) or selecting a cable TV operator (Lima & Fernandes, 2015). Experience services, in contrast to search services cannot be evaluated objectively as users must engage with them to evaluate them (which is subjective) (Xiao & Benbasat, 2007). Some examples of experience services are music (Chen & Chellappa, 2009), healthcare (Korachais et al., 2019) and hotels and restaurants (Zhang, Sun et al., 2014).

Search services "are more standardized and less personalized than experience or credence services … [and hence] pre-purchase judgments [of search services] are easier" (Chocarro et al., 2021, p.8). The perceived risk associated with a purchase is influenced by difficulty in judging the attributes; that is, "the harder it is to judge the attributes, the higher the perceived risk attached to the purchase." (Chocarro et al., 2021, p.8). Thus, the perceived risk in the purchase of experience services is higher than in the purchase of search services (Zhang, Wang, Wu et al., 2021). Hence for this study, we selected experience services as our context, and we selected restaurant reviews for data collection as restaurants are one of the most common services that are being used in daily life.

3.1.2 Subjects

Empirical test of our stated hypothesis can be based on students' sampling or through crowdsourcing such as Amazon MTurk. External validity is the most common shortcoming of student samples (Peterson & Merunka, 2014) as Kees at al. (2017) contended "findings from studies using student samples are limited in that these younger, geographically constrained, and relatively well-educated participants may not always be generalizable to broader populations of nonstudent adults" (p.142). However, MTurk enables collecting data from participants with various demographics hence it mitigates some external validity issues related to the student participants (e.g., limited age, income, and education) (Kees at al., 2017). In this research we use both student sampling as well as MTurk sampling to increase the generalizability of our findings. The complimentary characteristics of the two is "convenience sampling". Convenience samples "consist of consumers who are easily accessible rather than consumers who are randomly selected from the entire population of interest" (Kees at al., 2017, p.142). Therefore, many people of a target demographic have no chance of being chosen in a convenience sample, raising concerns about the convenience sample's representability of the target population (Kees at al., 2017). Nonetheless, extant literature in online review has adopted either student sampling or MTurk sampling. We have gone further to strengthen our findings by means of data triangulation (Hales, 2010, p.14) using both types of sampling.

3.1.2.1 Online Crowdsourcing Markets (OCMs)

An Online Crowdsourcing Market (OCM) is an "internet-based participant recruitment resource, which facilitates the distribution, completion and retrieval of survey responses" (Soror et al., 2015). Crowdsourcing has become more common in several fields in recent years (Fang & Chen, 2022) and is used in a variety of tasks, such as capturing new product ideas and innovations (Bayus, 2013), improving image search (Yan et al., 2010), healthcare applications (Hill et al., 2013), and processing social media data (Archak et al., 2011). OCMs are also utilized for detailed product reviews and experimental surveys (Steelman et al., 2014). Workers are paid a predetermined sum of money for successfully completing the given task, and the payment is processed through OCM payment mechanisms. (Steelman et al., 2014).

OCMs offer special advantages to researchers as there is a wide range of subjects in OCMs in terms of age, occupation, and culture, and compared to the student participants, they are generally available to participate in research studies (Steelman et al., 2014). Furthermore, researchers can restrict the participants to those who are qualified according to the research study requirements (Steelman et al., 2014). In all, OCMs are beneficial not only because they provide a wide range of participants with various levels of expertise and demographics but also at a low cost (Brynjolfsson et al., 2016)

3.1.2.2 Amazon Mechanical Turk (MTurk)

MTurk is one of the online crowdsourcing markets where "employers post outsourced tasks for an undefined, anonymous network of labourers to perform and receive compensation for their contributions. On Amazon's Mechanical Turk, registered users (called Workers [Turkers]) participate in tasks [called HITs, Human Intelligence Tasks] issued by individual employers (Requesters) that solicit the work." (Steelman et al., 2014, p.357). MTurk enables the researchers to assign tasks randomly to Turkers (Brynjolfsson et al., 2016), and once the task is completed upon the approval of the researcher, Turkers are paid automatically using the MTurk built-in system. MTurk is one of the popular OCMs used by academics: around "15,000 papers containing the phrase 'Mechanical Turk' were published between 2006 and 2014" on Google Scholar (Chandler & Shapiro, 2016, p.55) and in the last 6 years more than 340,000 MTurk studies were conducted by around 10000 researchers (Hauser et al., 2022). Although evidence suggests that participants (e.g., subjects recruited through MTurk, campus, and community) engage in unfavourable behaviours (e.g., cross-talking - talking with a participant who previously completed the task), compared to traditional studies, MTurk provides a large sample size which increases the statistical power of a research study (Necka et al.,2016). Additionally, it has been demonstrated that data gathered on MTurk for numerous psychological activities, including cognitive, social, and judgement and decision-making tasks, are often comparable to data gathered in a laboratory setting (Necka et al.,2016; Paolacci et al., 2010) and studies suggest that are motivated to provide high quality responses (Woo et al., 2015).

3.1.2.3 Suitability of MTurk for Data Collection

While collecting data using MTurk provides flexibility to the researchers, the first question that should be answered is the suitability of the MTurk data for the research question. Jia et al. (2017) contend that two dimensions of the research question should be considered in the assessments of using MTurk data: Generalizing vs. Contextualizing Study and Diverse vs. Shared

Cognition. These dimensions are shown in Table 5, and the use of MTurk data is most appropriate when the research question is a generalizing study with diverse cognition.

	Shared Cognition	**Diverse Cognition**
Generalizing Study	Cautioned	Appropriate
Contextualizing Study	Not Recommended	Cautioned

Table 5 Determination of the suitability of the MTurk data for a research question (Adapted from Jia et al. (2017))

Generalizing vs. Contextualizing Study: The spectrum of research questions varies from "generalizing" to "contextualizing" (Jia et al., 2017). In the case of a generalizing study for investigating generic attitudes and behaviours the use of MTurk participants would be appropriate, however in the case of a contextualizing study where the participants' identities are important their use may invalidate the findings (Jia et al., 2017).

For example, if a researcher is looking into a cross-cultural phenomenon, such as social adoption and usage of social media, then MTurk is a valid data source (Jia et al., 2017). However, if the study is specific to a situation, such as IT employee burnout in Silicon Valley firms, then it is crucial for participants to be a member of that specific situation (Jia et al., 2017). In these types of cases, the use of MTurkers should be avoided since their identity cannot be verified and hence can invalidate the results (Jia et al., 2017).

Diverse vs. Shared Cognition: Another dimension of a research question is the experiences of the individuals. Zhu et al. (2015) contend that the experience spectrum ranges from "diverse" to "shared." For example, if the research question is about understanding how technology adoption is affected by individual perceptions in general, then it is desirable to have participants from various backgrounds and experiences (Jia et al., 2017). On the other hand, if the research question requires some common experience, such as the employees' perceptions of a team environment (Jia & Reich, 2013), then a shared background for the study participants is required (Jia et al., 2017).[1]

Since our study is related to online reviews in the service industry (i.e., restaurants), our study focus is highly generic (i.e., consumers of the service industry) and doesn't require a shared cognition; hence, we contend that collecting data using MTurk is appropriate for our study. This is in line with the prior research which identified MTurk as suitable for conducting online

[1] "Shared cognition is the collective cognitive activity from individual group members where the collective activity has an impact on the overall group goals and activities" (Razzouk & Johnson, 2012, p.3056). "Knowledge possessed by effective teams has been referred to as shared knowledge, shared mental model, team knowledge, and shared understanding" (Cobb et al., 2014, p.17). Shared cognition includes "the knowledge that team members hold, which enables them to form accurate explanations and expectations for the task and in turn to coordinate their actions and adapt their behavior to demands of the task and other team members" (Razzouk & Johnson, 2012, p. 3056).

experiments (Han, 2021; Steelman et al., 2014; Mason & Suri, 2012; Paolacci et al., 2010). We would like to also state that "MTurk provides a larger and more diverse sample [compared to the college-student samples and community samples and] MTurk' subjects are more representative of the general population" (Han, 2021, p.262).

3.1.2.4 Challenges (and mitigations) of using MTurk

Although MTurk enables researchers to collect data from a large population at a low cost, it also requires a more thorough examination and validation of responses, e.g., detecting lower attentive participants or bot like - automated - responses (Steelman et al., 2014). Also, multiple responses from the same participants need to be scrutinized since it is advised for researchers to filter out workers who might have participated in an earlier version of a survey or pilot study to reduce bias (Steelman et al., 2014). Finally, to mitigate reliability and quality issues, it is recommended to recruit workers with high reputation ratings for their previous quality work (Steelman et al., 2014).

To mitigate the issues identified above, we recruited MTurk participants using the CloudResearch Mechanical Turk (MTurk) Toolkit. CloudResearch is an international online recruitment platform that is connected to more than 50 million research participants worldwide. CloudResearch uses MTurk's application programming interface (API) for connection, and it provides data quality checks otherwise not available to MTurk. Such quality checks include excluding participants who fail the attention checks or who provide bot-like responses. Furthermore, CloudResearch also has the capability of blocking duplicate IPs or suspicious geolocations (i.e., locations that create high web traffic) to increase the data quality of responses.

In addition, it has been suggested that compared to the other data collection tools (e.g., Prolific) CloudResearch data is superior (Litman et al., 2021).

In addition to the above challenges, *self-selection bias, self-misrepresentation, inconsistent English language fluency, MTurker non-naivete, MTurker inattention, high attrition rates, perceived researcher unfairness, MTurker social desirability bias,* and *Vulnerability to web robots (or "bots")* need to be also addressed while using MTurk for data collection (Aguinis et al., 2021). We briefly describe these challenges and provide a mitigation strategy for them, as suggested by Aguinis et al. (2021).

Self-selection bias: Self-selection bias occurs "when survey respondents are allowed to decide entirely for themselves whether or not they want to participate in a survey", resulting in biased data in the sense that the survey participants do not represent the target population (Lavrakas, 2008, p. 809)

Our design: One of the suggestions related to the self-selection bias is to assess the alignment between the target population of the study and the MTurkers (Aguinis et al., 2021). In our study, we are interested in the consumers' experiences in the service industry (i.e., restaurants), which don't require specialized knowledge. Hence, we do not expect any misalignment between the target population (i.e., restaurants' customers) and the MTurkers. However, as Hong et al. (2016) contended, culture is one of the factors that affect consumer behaviour. To mitigate the possible effects of culture on the restaurant evaluations, we hired only US and Canadian residents for the study. This is controlled by both verifying the country of the participants' IP addresses and the CloudResearch tools, as CloudResearch provides researchers to set a restriction on the survey participants' country location via the "Verify Worker Country Location" option.

Self-misrepresentation: In order to participate in a survey, MTurkers may misrepresent their characteristics such as income, age, education or gender (Aguinis et al., 2021).

Our design: Since we expect "diverse cognition" in our study, we only imposed an age limit (18 +) and US/Canadian residency on the study participants. Thus, self-misrepresentation is mitigated by not restricting the participants by their characteristics. Similar to Reich et al. (2023), we also utilized CloudResearch's 'Approved Participants' feature for US participants "to ensure the participation of only high-quality participants who have passed CloudResearch's attention and engagement measures"(p. 294).

Inconsistent English language fluency: Survey participants' level of fluency in English can affect the survey results since it can affect the participants' interpretations of the measures and instructions (Aguinis et al., 2021).

Our design: Since we restricted the participants to be US and Canadian Residents, we do not expect any inconsistencies in the level of English language of the MTurk participants.

MTurker non-naivete: Participants' previous survey experiences can impact the study's findings since some participants might have been exposed to the same stimuli before, and they are more knowledgeable about the experiment compared to novice participants (Aguinis et al., 2021).

Our design: Since our study requires no specialization and MTurkers will write a review about their own restaurant experiences (or rate a previously written review), we expect the effect of non-naivete to be minimal.

MTurker inattention: Approximately 15% of MTurk responses fail to comply with survey checks since the participants, to maximize the monetary returns, fail to pay enough attention to the survey instructions (Aguinis et al., 2021).

Our design: As stated before, CloudResearch's data quality filters enable high-quality data collection (Litman et al., 2021). We used the "Approved Participants" option whenever possible to filter out the low-quality MTurk participants. Also, we imposed MTurkers to provide details about their restaurant visit and write a review with at least 100 words for their restaurant experience. Hence we expect the effect of inattention to be minimal for our study.

High attrition rates: Due to its online nature, some Mturk participants may not complete the survey, and the percentage can exceed 30% (Aguinis et al., 2021).

Our design: To mitigate the attrition rates, we paid US$2 to the participants who completed the survey successfully. Participants were able to see the allotted compensation (US$2) before signing up for the task. As the study was estimated to take approximately 12 minutes to complete, "this amount is comparable to MTurk studies [e.g., Martin & Nissenbaum, 2017; Elias et al., 2016] of similar length" (Clauss & Bardeen, 2022, p.520).

Perceived researcher unfairness: In case participants feel that the researcher is unfair, they can boycott the researcher's next studies (Aguinis et al., 2021).

Our design: To mitigate such misunderstandings and to make sure that the instructions are clear, we conducted a pilot study with 8 participants to ensure the clarity of the instructions in the experiment. For the controlled experiment, we included the details about the compensation rules (e.g., refuse to pay conditions), the researcher's contact details, and McMaster University's ethics board's contact details provided in the consent form. We received two emails from the survey participants about completing the survey, and we responded to them within one day of receipt of the emails.

MTurker social desirability bias: Due to monetary compensation, MTurkers are "more likely to provide socially desirable responses than student samples" (Aguinis et al., 2021, p.827).

Our design: Considering that we asked participants to write reviews (or rate a written review) for restaurant experiences, we don't expect to observe social desirability bias in our study. To further mitigate this concern, we will use students as subjects for the second phase of our study.

Vulnerability to web robots (or "bots"): Since MTurk provides an online survey, it may be the case that some of the MTurkers use programs called "bots" that can answer the survey questions randomly (Aguinis et al., 2021).

Our design: Although participants can use bots, we believe the use of bots is minimal, if not impossible, since:

a) We implemented a "CAPTCHA" verification (i.e., we asked for the sum of two integer numbers and allowed the participants to start the survey only if they could answer the result correctly) at the start of the study to thwart web robots
b) The participants will be asked to write reviews in addition to completing a short survey. We disabled copy-paste on the survey form, and we also counted the number of words in the review.

3.1.2.5 Student Participants

Recent studies compared the performances of the online and student participants, and they found that online participants are at least as good as student participants (Dalton, 2021) in terms of "effort, honesty, self-reported numeracy, and analytical ability" (Dalton, 2021, p.35). Additionally, researchers have found that the data from MTurk is comparable to that collected from college students or marketing research firms (Fowler et al., 2022) in terms of "measures of engagement, indices of test-retest reliability and internal consistency, and measures of criterion validity" (Fowler et al., 2022, Reasons to be Enthusiastic About Crowdsourced Research section, para. 2). For the US, it was also found that the MTurk data samples "produced models that lead to similar statistical conclusions as both U.S. students and U.S. consumer panels" (Steelman, 2014, p.355).

Although MTurk produces comparable results to the student participants, it has been cautioned to use MTurk excessively to recruit participants from MTurk as very little is known about the MTurk participants (Krupnikov & Levine, 2014). That is, as stated by the self-selection bias, MTurk subjects participate in a research study voluntarily rather than being selected randomly from a pre-specified population (Mullinix et al., 2015).

Thus, to mitigate possible concerns about the MTurk data and to strengthen conclusions about our findings, we also recruited student participants for our study. By using both MTurk and student participants, this study is also able to assess whether student participants and MTurk participants show differences in their online review evaluations.

University students are suitable for our research as they are considered one of the consumer profiles of the restaurants that use online reviews (Souki et al., 2022). In addition, university students are very similar in terms of "age, intelligence and income so this similarity can reduce the potential effects of these potential covariates in the results." (Liu et al., 2012, p.928). Furthermore, studies showed college students' online behaviour is similar to that of the general population (Wang et al., 2020). Therefore, we contend that university students are appropriate subjects for our study.

3.1.3 Pilot Study

Prior to data collection, a pilot study was conducted with 8 Ph.D. students to ensure the instructions and measurement instruments are clear, and online survey works as expected. After the experiment, participants' feedback confirmed that the online surveys are clear and there is no technical difficulty in the online surveys' platform. After the pilot study, we conducted the main controlled laboratory experiment using the following procedures.

3.2 Experimental Procedures

The controlled laboratory experiment was conducted using the Qualtrics online survey platform, where subjects can fill out the survey form using a web browser at a suitable time of their choice.

3.2.1 Experiment, Phase 1:

This experiment was conducted to collect online reviews about restaurants. As part of a simulated service provider feedback scenario, each participant was asked to think about a specific restaurant visit and write a review about it (See APPENDIX – A Survey Form for Online Review Writers).

3.2.1.1 Subjects

The research was conducted by using two groups of participants: 256 participants from Amazon Mechanical Turk (MTurk) and 327 student participants from McMaster University enrolled in the Commerce program.

3.2.1.1.1 MTurk Participants

Data collection took place between the 20th of July 2021 - the 9th of August 2021, and to mitigate possible cultural factors in the study, we recruited participants from the US and Canada.

The number of participants in this part of the study was 256 (81 from Canada and 175 from the US). Participants were paid US$2 via CloudResearch upon successfully completing the survey.

3.2.1.1.2 Procedure

1. Consent Form: As per McMaster research ethics board (MREB) requirements, before the start of the experiment, we asked all the subjects to accept/reject the consent form explaining the nature of the study. After accepting the consent form, the subject was allowed to take the survey.

2. Instruction Form: A short instruction form about the contents of the survey was presented to the participants. The subjects were allowed to continue to survey by pressing Next.

3. Preliminary Questions: A set of screening questions was presented to the subjects. The subject was allowed to take the survey if they held all of the following conditions:

    a. Having an age of 18 years or older
    b. Being a Canadian or US resident: this is to mitigate any effects of culture on the restaurant evaluations of review writers in North America. A Canadian/US "restaurant setting may not be applicable to other countries or cultures [as] consumers from other cultures may behave differently on their online review behavior" (Li, Zhang, et al., 2019). For instance, Hong et al. (2016) report that compared to consumers from collectivist cultures, consumers from individualist

cultures are more likely to write reviews and include more emotional expressions in their reviews.

c. Have been to a mid-price range (according to TripAdvisor, it is $$-$$$) restaurant in the last six weeks. Price is one of the features consumers use to evaluate a product/service (Chua et al., 2020) as Lin et al. (2020) contended consumers are willing to pay more when they consider the food item has high quality, is locally sourced or is labelled as "safe" according to the food safety standards. In a restaurant setting, price provides information about the perceived quality of the service or food (Kim et al., 2022), and consumers use price as one of the factors in the restaurant assessment (Pantelidis, 2010). For instance, when a restaurant's price and ambiance are considered together, a reviewer's anticipation of ambiance is strongly correlated with price (Luo et al., 2020). To mitigate price-related differences in the restaurant, we set the price range to a mid-price range for all participants. Tripadvisor provides three price ranges for restaurants, and they are indicated by $ (Cheap Eats), $$-$$$ (Mid-Range) and $$$$ (Fine Dining). For the purpose of this study, we selected mid-price range ($$-$$$) restaurants as our context, similar to Chow et al. (2007), and Kim and Velthuis (2021). We selected mid-price restaurants as consumers are more like to check online reviews for such restaurants (Kim & Velthuis, 2021). We excluded low-price restaurants since they are less likely to refer to online reviews for such restaurants (Kim & Velthuis, 2021) as "cheap establishments remain attractive regardless of their online evaluation" (Beuscart et al., 2016, p. 462). We also excluded high-price restaurants as high-priced restaurants "have sufficient reputation capital to not economically suffer or benefit from online reviews" (Beuscart et al., 2016, pp. 463). In addition, consumers of high-price restaurants are expected to use other resources, such as the Michelin guide, which has more expertise and legitimacy (Kim & Velthuis, 2021). While some review platforms allow anyone to post a review, some review platforms require a user to post a review only if they had the service in the last six months. (Nam et al., 2020). In this study, six weeks are selected to ensure that the participants' restaurant experience is recent: we expect participants can provide more and accurate details about their restaurant visits if they visited a restaurant recently.

d. Have posted online reviews in the past: this is to ensure that the participants are familiar with and knowledgeable about online reviews.

4. Subjects were asked to provide demographic information as follows: *age*, *gender* and *education level*. Personal characteristics can affect consumer evaluations as Sharma et al. (2012) found that male and older customers are more demanding and have higher expectations than female and younger customers. Education level is also another factor in the service evaluation. For hotel restaurant complaints, Heung and Lam (2003) found that 67.2 percent of the complainers have at least university education, and close to 45 percent of the non-complainers have only primary education.

5. Then the subjects were asked to provide

a. The name of the restaurant as shown on the TripAdvisor website, the state/province and city of the restaurant and the date of the restaurant visit: This information is required to make sure that the particular restaurant is within a mid-price range restaurant and to make sure that the participant has been to the particular restaurant within the last six weeks as stated in the preliminary questions.

b. Whether they have written about this particular restaurant before. We expect participants who wrote a review for this restaurant to recall their restaurant visit experience better compared to those who haven't written a review. However, this data is collected for reporting purposes, and no response is removed if a participant hasn't written a review for this restaurant before.

c. The subjects were then asked to write a review about the particular of their restaurant experiences. "A minimum text length is required to reach stable text coverage" (Zhang, Wang, Chen et al., 2021, p. 8), "to measure the intelligibility of reading materials" (Chujo & Utiyama, 2005, p.1). For psychological text analysis, a minimum number of words are suggested by the experts (Alzate et al., 2022): while Boyd (2017) recommended 25–50 words per text, Chung and Pennebaker (2019) and Tang and Guo (2015) recommended 100 words per text. Hence, we set the minimum text length to 100 words. This was ensured by counting the number of words in the review. Furthermore, to make sure that the participants provided their own experience rather than copy/paste from another source (e.g., a review website), we disabled the copy-paste facility for review writing.

d. Next, the subjects were asked to assign a star rating to their restaurant experience using a 5 point-Likert scale where 1 star represents an extremely negative experience, and 5 stars represent an extremely positive experience of the service. Star rating represents the online review senders' attitudes towards their restaurant experience that is operationalized on online review sites (e.g., Yelp.com).

e. Next, the subjects were asked to provide their positive and negative evaluations of the particular restaurant. These evaluations are used in our analyses within the context of the "Evaluative Space Model" (ESM) (Cacioppo et al., 1997). ESM contends that an attitude is formulated as a combination of positive and negative evaluations. Positive and negative evaluations are measured using the following two questions (adopted from Snyder and Tormala (2017)):

    i. Considering only your POSITIVE thoughts and feelings about your restaurant visit and ignoring the negative ones, how positive would you say your positive thoughts and feelings are? (Likert scale between 0-9) **[Positive Evaluation]**

    ii. Considering only your NEGATIVE thoughts and feelings about your restaurant visit and ignoring the positive ones, how negative would you say your negative thoughts and feelings are? (Likert scale between 0-9) **[Negative Evaluation]**

These positive and negative evaluations are used to calculate the objective ambivalence of the subjects towards the restaurant using the SIM (Similarity-Intensity model) (Thompson et al., 1995):

$$Objective\ Ambivalence = \frac{Positive\ Evaluation\ +\ Negative\ Evaluation}{2} - abs\ (Positive\ Evaluation - \ Negative\ Evaluation)$$

3.2.1.1.3 Student Participants

The subjects in the experiment were 327 undergraduate students enrolled in a second-year commerce course entitled "Information Systems in Business" at the DeGroote School of Business in the Fall 2021 term. Students were compensated with a 2% bonus mark towards their final course grade for their participation in the experiment.

3.2.1.1.4 Procedure

The experimental procedure for the student participants was exactly the same as for the MTurk participants, except that there were no preliminary questions (i.e., questions stated under the procedure section: section 1.2.1.1.2 - item 3) as required by the McMaster research ethics board (MREB), and the subjects were asked to provide their personal details (i.e., Student ID, First and Last name) for bonus mark compensation.

3.2.2 Experiment, Phase 2:

This experiment was designed to test the ambivalence evaluations of online review readers in an experimental environment. As part of a simulated service provider feedback scenario, each participant was asked to read an online review about a restaurant and then provide their evaluations of the restaurant in the review (See APPENDIX – B Survey Form for Online Review Readers for details).

3.2.2.1 Subjects

The research was conducted by using two groups of participants: 252 participants from Amazon Mechanical Turk (MTurk) and 657 student participants from McMaster University enrolled in the Commerce program.

3.2.2.1.1 MTurk Participants

Similar to the Experiment in Phase 1, participant recruitment was done using the CloudResearch Mechanical Turk (MTurk) Toolkit. To mitigate cultural factors in the study, we recruited the participants from Canada and US only. The set of participants is different from the participants in Phase 1 as CloudResearch provides functionality to exclude the previous participants from consecutive experiments. The number of participants in this part of the study is 252. Participants were paid US$2 via CloudResearch upon successfully completing the survey.

3.2.2.1.2 Procedure

1. Consent Form: As per McMaster research ethics board (MREB) requirements, before the start of the experiment, we asked all the subjects to accept/reject the consent form explaining the nature of the study. After accepting the consent form, the subject was allowed to take the survey.

2. Instruction Form: A short instruction form about the contents of the survey was presented to the participants. The subjects were allowed to continue to survey by pressing Next.

3. Preliminary Questions: A set of screening questions was presented to the subjects. The subject was allowed to take the survey if they held all of the following conditions:

a. Having an age of 18 years or older

b. Being a Canadian or US resident: Similar to review writers in Phase - 1, this is to mitigate any effects of culture on the restaurant evaluations of review readers.

c. Have read and rated online reviews about restaurants in the past: This is to ensure that the participants are familiar with and knowledgeable about online reviews.

4. Subjects were asked to complete an online questionnaire to collect their demographic information: *age*, *gender* and *education level*. This was similar to review writers in phase – 1.

5. Then, subjects were randomly shown a review from the pools of reviews collected in Experiment, Phase 1, as follows:

   The following Review is written by a **very knowledgeable reviewer** for a **mid-ranged priced** restaurant:

   <Review>

6. Similar to phase-1, the subjects were then asked to provide the following:

   a. Their positive and negative evaluations of the particular restaurant using the following two questions:

      i. Considering only your POSITIVE thoughts and feelings about this RESTAURANT and ignoring the negative ones, how positive would you say your positive thoughts and feelings are? (Likert scale between 0-9) **[Positive Evaluation]**

ii. Considering only your NEGATIVE thoughts and feelings about this RESTAURANT and ignoring the positive ones, how negative would you say your negative thoughts and feelings are? (Likert scale between 0-9) **[Negative Evaluation]**

These positive and negative evaluations are used to calculate the objective ambivalence of the subjects using the SIM (Similarity-Intensity model) (Thompson et al., 1995):

$$Objective\ Ambivalence = \frac{Positive\ Evaluation\ +\ Negative\ Evaluation}{2} - abs\ (Positive\ Evaluation - \ Negative\ Evaluation)$$

b. Next, the subjects were asked to indicate their perception of the overall rating of the restaurant in the Review using a 5 point-Likert scale where 1 star represents an extremely negative evaluation, and 5 stars represent an extremely positive evaluation. The subjects were asked the following question to rate the restaurant:

Please indicate your perception about the overall rating of this RESTAURANT

d) Next, the subjects were asked to indicate their perception of the review's helpfulness using three questions measured on a nine-point scale (adapted from Yin et al. (2014)). The subjects were asked the following questions to measure review helpfulness:

Using the scales below, how would you describe the above consumer review?

i. not at all helpful/very helpful (1 not at all helpful, 9 very helpful) **[Helpful Evaluation]**

ii. not at all useful/very useful (1 not at all useful, 9 very useful) **[Useful Evaluation]**

iii. not at all informative/very informative (1 not at all informative, 9 very informative) **[Informative Evaluation]**

Review helpfulness is calculated by averaging the above items' scores:

$$Review\ Helpfulness = \frac{Helpful\ Evaluation\ +\ Useful\ Evaluation + Informative\ Evaluation}{3}$$

3.2.2.1.3 Student Participants

For this part of the experiment, student participants were asked to evaluate the online reviews of the restaurants:

a) written by the MTurk participants and collected in Experiment, Phase 1. The subjects in the experiment were 333 undergraduate students

b) written by the student participants and collected in Experiment, Phase 1. The subjects in the experiment were 324 undergraduate students

All of the students are enrolled in a second-year commerce course entitled "Information Systems in Business" at the DeGroote School of Business in the Fall 2021 term. Students were compensated with a 2% bonus mark towards their final course grade for their participation in the experiment.

3.2.2.1.4 Procedure

The experimental procedure for the student participants was the same as for the MTurk participants, except that there were no preliminary questions (as required by the McMaster research ethics board (MREB)), and the subjects were asked to provide their personal details (i.e., Student ID, First and Last name) for bonus mark compensation.

3.2.2.1.5 Control Variables

We controlled the following factors in the experiment:

1. Reviewer-related factors: In an online review platform, online review readers can be affected by the reviewer's expertise level (e.g., the "Elite" badge on Yelp.com or the "Top 10,000 Reviewer" badge on Amazon.com), reviewer's trustworthiness (e.g., the number of reviews submitted by a sender (Filieri, 2016) or the number of followers of a reviewer (Banerjee et al., 2017) or their attractiveness (e.g., the number of the sender's online friends (Zhu et al., 2014)). Hence, review expertise is controlled in the experiment by stating that "the provided Review is written by a very knowledgeable reviewer".
2. Restaurant-related factors:

a. Restaurant Price: As stated previously, price provides information about the perceived quality of the service or food in a restaurant setting (Kim et al., 2022); hence, consumers set expectations based on the price resulting in them evaluating the service quality differently than the pre-consumption stage (Kim et al., 2022). Also, restaurant price can be an important consideration for customers in choosing a particular restaurant (Chow et al., 2007). Hence restaurant price is controlled in the experiment by stating that the provided "review is written for a mid-ranged priced restaurant".

b. Restaurant Name: Some of the reviews written by students contained the name of the restaurant. Since "university students live in a close-knit community" (Ab Rahman, 2005, p.596), and they may be knowledgeable about the same restaurants. If a student is knowledgeable about the restaurant stated in the review, they may provide their evaluations based on their own experiences rather than the statements provided in the review. Hence, to mitigate any bias related to the restaurant, we removed the name of the restaurant in the review before showing the reviews to the student review readers.

Table 6 presents the variables controlled in the study.

	Variable	Explanation
Review Writer	Expertise	To mitigate any biases related to reviewer expertise, we stated to the review readers that "the provided Review is written by a very knowledgeable reviewer".
Restaurant	Price	Since price could affect restaurant evaluations, we controlled the restaurant price by stating to the review readers that the provided "review is written for a mid-ranged priced restaurant".
	Name	To mitigate any bias related to the restaurant, we removed the name of the restaurant in the review before showing the reviews to the student review readers.
Review	Review Length	We asked the participants to write at least 100 words for the reviews for proper text analysis.
Review Reader/Writer	Location	To mitigate any cultural effects on the study, we recruited participants only from the US and Canada.

Table 6 Variables controlled in the study

3.3 Text Mining

Text mining "also called text analysis, text analytics, text data mining, automatic text analysis, and computer-based text analysis is the analysis of text data in order to discover hidden patterns, traits, and relationships" (Tang & Guo, 2015, p.68). Some of the tasks used in text mining are information extraction, text summarization, text categorization and sentiment analysis (Jo, 2019; Weiss et al., 2015; Boiy & Moens,2009). Our focus in this research is on using text mining for sentiment analysis. Sentiment analysis, or opinion mining, "is an active area of study in the field of natural language processing that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions via the computational treatment of subjectivity in a text" (Gray et al., 2023, p.548). For example, Tang & Guo (2015) used text mining to show that the "linguistic indicators generated by text analysis are predictive of eWOM communicators' [i.e., online review writers'] attitudes toward a product or service"(p., 67). To further strengthen our findings, we also utilized text mining in this study.

There are several techniques to assess review texts, such as topic mining, semantic text analysis, word embedding, and opinion mining. However, since the objective of this research is to test the stated hypothesis "the anticipatory ambivalence of the review readers based on a two-sided review content for a focal service is higher than the ambivalent attitude of the writer of the review who has already experienced the focal service" both the positive/negative evaluations of review writers' and review readers' are needed to assess their ambivalence values. Next, we provide some examples related to topic mining, semantic text analysis and opinion mining (sentiment analysis) and then provide our rationale for using the sentiment analysis in the study.

Topic mining: "Topic mining is a statistical method used to discover latent topics in a series of documents" (Zhang et al., 2020, p. 64820).

For example, let's assume that the following documents are provided for topic modelling.[2]

Document 1: We watch a lot of videos on YouTube.

Document 2: YouTube videos are very informative.

Document 3: Reading a technical blog makes me understand things easily.

Document 4: I prefer blogs to YouTube videos.

Topic modelling then can provide the following topics:

Topic 1: associated words -- videos and YouTube

[2] These example are taken from https://www.analyticsvidhya.com/blog/2021/07/topic-modelling-with-lda-a-hands-on-introduction/

Topic 2: associated words – blogs and YouTube

Based on the generated topics, the algorithm automatically assigns the topics to each of the above documents. For example, Document 1 can be assigned to Topic 1 as it covers both "Youtube" and "videos." The SIM (Similarity Intensity Model) used for ambivalence calculation requires both positive and negative evaluation about an entity (e.g., restaurant). Since topic modelling doesn't provide any information about the text's positivity and negativity, we contended that it is not a suitable technique for the current study.

Semantic Text Analysis: "Semantic text analysis is representation of information based on meaningful relationships of a written text, structured as a network of words, cognitively associated with one another" (Bayrakdar, 2020, p.11). One of the tasks involved in semantic analysis is relationship extracting. It involves identifying various entities in the sentence and then extracting the relationships between them. For example, for the sentence "Casa Madera restaurant is a unique experience in the city of Toronto. Shout out to Roberts for the great service!"[3], the semantic analysis provides the extraction of the following entities (e.g., people, organizations, places) and their relationships:

Casa Madera restaurant is a unique experience in the city of Toronto.

[Organization] [Location]

Shout out to Roberts for the great service!

[Person]

[3] This is a part of review taken from TripAdvisor about the Casa Madera restaurant.

Semantic text analysis also doesn't provide any information about text positivity and negativity that is required for computing ambivalence as modelled by the SIM. Hence, we contend that it is not a proper technique for this study.

Opinion Mining: "Opinion mining [also known as sentiment analysis] is a Natural Language Processing task that aims to determine a person's attitude by identifying and extracting information. The major task in opinion mining is to classify the polarity of a review at sentence level, whether the expressed opinion is positive or negative"(Hasan et al., 2015, p.511).

For example, opinion mining (sentiment analysis) provides the following (overall) sentiments for the given sentences:

Shout out to Roberts for the great service! – Positive

A hype place, completely overpriced. – Negative

The restaurant is located in downtown Toronto. – Neutral

We contended that opinion mining is the suitable technique for this study since the SIM model used for ambivalence calculation requires both positive and negative evaluations and opinion mining (sentiment analysis) provides values for both text positivity and negativity,

Tools available for sentiment analysis of the reviews include: Google Cloud Natural Language API, IBM Watson Natural Language Understanding, and VADER (Valence Aware Dictionary and sEntiment Reasoner). VADER assesses both the positive and negative sentiment scores in addition to the overall sentiment scores. While the other tools only assess the overall sentiment of the review (i.e., whether the count of overall sentiment of the review is positive or negative). To that end, we used VADER because anticipatory ambivalence assessment, for testing the stated hypothesis, requires assessment of positive and negative sentiments separately to measure the ambivalence. This is necessary, because Similarity Intensity Model (Thompson et al., 1995) (P + N)/2- abs (P - N) utilized in this research take into account the different effects of positive and negative sentiments in terms of positivity offset and negativity bias. VADER is one of the most popular sentiment analysis tools (Smirnov & Hsieh, 2022).

VADER is a dictionary and rule-based tool developed by Hutto and Gilbert (2014) that enables assessing a sentence's sentiment and intensity. VADER uses a dictionary developed from commonly used word banks such as LIWC (Linguistic Inquiry Word Count), ANEW (Affective Norms for English Words) and GI (General Inquirer) and extended the dictionary by including emotions, acronyms and slangs used in online communities (Hutto & Gilbert, 2014). The tool uses internal scores and rules to calculate the sentiments of texts. It was used in previous studies (Borg & Boldt, 2020; Deng, 2020; Zhang et al., 2019; Park & Seo, 2018), and according to Hutto and Gilbert (2014), it performs better compared to other sentiment analysis tools within social media texts.

4 Analysis

To test our hypothesis, we utilized t-tests. As stated in the previous section, we used four scenarios to assess the transfer of knowledge (emotions/attitudes) from the review writers to review readers. For each scenario, we used the t-test to test the stated hypothesis: if there is a significant difference between the review readers' and review writers' emotions/attitudes towards the restaurant stated in the review. Specifically, we applied the following two steps to test the hypothesis:

1. Apply independent sample t-tests to test the mean differences in the review readers' and review writers' emotions/attitudes to test the hypothesis.
2. Employ text mining to extract review readers' and review writers' emotions/attitudes from the text and then conduct the independent sample t-test to test the hypothesis using the values extracted from the text.

To that end, we first applied preliminary data-cleaning activities to ensure the data is suitable for analysis. We elaborate on the results of these analyses as follows.

4.1 Data Cleaning

As stated before, during the controlled laboratory experiment, we collected data from Amazon Turk participants in Canada and the United States and from student participants enrolled in the "Information Systems in Business" course at the DeGroote School of Business. We removed some of the data according to the following rules:

1. Less than 100 words for the text data: During the experiment, we asked participants to write an online review of at least 100 words (for review writers) or to provide their views on the provided review with at least 100 words (for review readers). We found that some 3

MTurk (2 Canadian and 1 US MTurk review writer) and 9 student review writers and 7 (3 for MTurk reviews and 4 for student reviews) student readers wrote less than 100 words. Therefore, we removed those responses from the dataset.

2. Missing MTurk ID: During the experiment, we asked the MTurk participants to provide their MTurk IDs to check if they are legitimate MTurk users. We compared the provided MTurk ID with the MTurk ID obtained via CloudResearch and removed 3(1 US MTurk writer's MTurk ID was not saved by CloudResearch – associated review readers' responses were also removed; 1 MTurk readers entered N/A as MTurk ID and 1 MTurk reader's MTurk ID is not saved by CloudResearch) responses where there was a mismatch between these two.
3. Participants' Country of Residence: To mitigate possible cultural effects in the restaurant evaluations, we only required participants from the US and Canada. We utilized CloudResearch's "Verify Worker Country Location" and the IP addresses of the participants to check if the participants were from the US or Canada. We found that 1 participant's IP address was outside of US or Canada; hence we removed those responses from the analysis.
4. For student responses we corrected the incorrect student IDs and removed 5 student (1 student for student review writer -associated readers' responses were removed as well-; 2 students for MTurk review readers and 2 students for student review readers) responses since their student IDs couldn't be found in the course list. Furthermore, we removed 4 (3 students for MTurk review readers and 1 student for student review readers) students' responses (in total 8 responses) since they completed the same survey more than once. One student completed the survey twice (one for MTurk review reader and one for student

review reader) hence those responses removed. For student review writers we noticed that 40 restaurants are less than mid-price ranged ($$-$$$) or price information is not available on TripAdvisor hence their responses are removed from the student review writer dataset. We also noticed that one student hadn’t written 100 words properly and three students' responses have minor errors hence their responses were removed from the student review writer's dataset. These 44 removed reviews are not shown to the student review readers. We noticed that two student review writers attempted the survey twice; although, they fully completed the survey only once, their responses and associated student readers' responses (three in total) are removed from the dataset. Furthermore, one student for MTurk review evaluation and four students for Student review evaluation attempted to complete the survey two times. Hence, their responses were also removed from the dataset.

5. Due to Qualtric's working mechanism, some reviews (obtained from review writers) weren't shown to the review readers, and some of the reviews were shown to more than one review reader. Hence, for our analysis, we only considered those reviews that are shown and evaluated by the review readers (removed 41 reviews written by review writers).

The application of the above steps resulted in the following usable responses[4]:

Scenarios	# of obs (reader)	# of obs (writer)
MTurk CA Reader - MTurk CA/US Writer	90	90
MTurk US Reader - MTurk CA/US Writer	157	135
Student Reader - MTurk CA/US Writer	319	228
Student Reader - Student Writer	309	238

Table 7 Number of observations for each scenario after data cleaning

Descriptive statistics for the collected data are provided in the appendix (see Appendix A)

4.2 Hypothesis Testing

To test our hypothesis, t-tests were conducted. We compared review writers' ambivalent attitudes towards the restaurant that they visited - and wrote a review - with the review readers' ambivalent attitudes towards the same restaurant stated in the review. We utilized the t-test for each aforementioned scenario by using the following two data values:

I. Using review readers' and writers' objective ambivalence values assessed directly from response to the questionnaire (See APPENDIX – A Survey Form for Online Review Writers and APPENDIX – B Survey Form for Online Review Readers).

[4] We noted that some of the MTurk review writers wrote their reviews for a restaurant which they visited more than six weeks ago. We haven't removed any data due to this.

II. Using review readers' and writers' objective ambivalence values obtained using text mining.

Before conducting the t-test we applied ANOVA to verify if there are any differences in the review readers' restaurant evaluations (objective ambivalence) for the 3 groups: MTurk Canadian review readers, MTurk US review readers and Student review readers. ANOVA analysis resulted that there is a statistically significant difference between the review reader groups' objective ambivalence evaluations with F= 21.40 and p-value<0.01. We checked the homogeneity of the variances of the groups using the Levene's test. The test results indicated that the groups' objective ambivalence values have equal variances (pvalue=0.06). As Sainani (2012) stated, one of the assumptions of linear models such as ANOVA is the normality of the variables. We used the Shapiro-Wilk test for normality assessment. The results revealed that the groups' objective ambivalence values are not normally distributed. However, scholars identified that ANOVA analysis is robust to non-normality (Cheng & Ku, 2009) Therefore we consider that ANOVA analysis is appropriate. Although ANOVA analysis revealed that the group means are different, it doesn't provide information about which group means are statistically different from each other. To identify the differences between the groups we utilized the Tukey's test. The results are presented in Table 8. As the results indicate there is a significant difference (at the 5% level) in the review readers' mean objective ambivalence values between MTurk CA readers (0.92) and MTurk US readers (-0.14), and MTurk US readers (-0.14) and Student readers (1.69). The difference between MTurk CA readers (0.92) and the Student readers (1.69) is significant at the 10% level.

Group 1	Group 1 Mean Objective Ambivalence	Group 1 # Obs	Group 2	Group 2 Mean Objective Ambivalence	Group 2 # Obs	Mean Difference (Group 2 Mean - Group 1 Mean)	p-adj.
MTurk CA Reader	0.92	90	Student Reader	1.69	628	0.77	0.08
MTurk CA Reader	0.92	90	MTurk US Reader	-0.14	157	-1.06	0.03
Student Reader	1.69	628	MTurk US Reader	-0.14	157	-1.83	<0.01

Table 8 Tukey's test results for review readers for the three groups

Furthermore, we also checked the differences in the review writers' objective ambivalence values for the 3 groups: MTurk Canadian review writers, MTurk US review writers and Student review writers. ANOVA analysis resulted that there is a statistically significant difference between the review reader groups' objective ambivalence evaluations with F= 10.28 and p-value<0.01. We checked the homogeneity of the variances of the groups using the Levene's test. The test results indicated that the groups' objective ambivalence values have equal variances (pvalue=0.48). Shapiro-Wilk test for normality assessment revealed that the groups' objective ambivalence values are not normally distributed. However, as previously stated ANOVA analysis is robust to non-normality (Cheng & Ku, 2009); hence, we consider that ANOVA analysis is appropriate. Similarly, we utilized the Tukey's test to identify the differences between the groups. The results are presented in Table 9. As the results indicate there is a significant difference (at the 5% level) in the review writers' mean objective ambivalence values between MTurk CA writers (-1.20) and Student writers (0.21) and MTurk US writers (-1.12) and Student writers (0.21). There is no significant difference between MTurk CA writers (-1.20) and the MTurk US writers (-1.12).

Group 1	Group 1 Mean Objective Ambivalence	Group 1 # Obs	Group 2	Group 2 Mean Objective Ambivalence	Group 2 # Obs	Mean Difference (Group 2 Mean - Group 1 Mean)	p-adj.
MTurk CA Writer	-1.20	59	Student Writer	0.21	238	1.42	<0.01
MTurk CA Writer	-1.20	59	MTurk US Writer	-1.12	148	0.08	0.98
Student Writer	0.21	238	MTurk US Writer	-1.12	148	-1.33	<0.01

Table 9 Tukey's test results for review writers for the three groups

As the above analyses indicate there is a significant difference in the review readers' objective ambivalence evaluations at the 10% level. Since review readers can read the online reviews written by individuals with various backgrounds (written by MTurks, students or professional travelers), we contend that it is review readers' assessment of the review that creates the difference between the review readers' and review writers' evaluation of the focal service stated in a review text. Considering that there is a significant difference in the Student review writers' and MTurk CA/US review writers' objective ambivalences we decided to go with the four scenarios depicted in Table 7: namely MTurk CA review readers assessing the reviews written by MTurk CA/US review writers (scenario 1), MTurk US review readers assessing the reviews written by MTurk CA/US review writers (scenario 2), Student review readers assessing the reviews written by MTurk CA/US review writers (scenario 3) and Student review readers assessing the reviews written by Student review writers (scenario 4). For each of three scenarios, t-test is applied to test our hypothesis that review readers are more ambivalent than review writers.

In order for t-test analysis to provide appropriate statistical results, required statistical power needs to be ensured. Generally, the statistical power of 0.8 is accepted as appropriate for statistical analysis (Myors et al., 2008). We used "G*Power 3.1" software (Faul et al., 2007) to calculate the minimum sample size required for the statistical power level of 0.80 at $\alpha = 0.05$ is satisfied for t-test analysis for medium effect size (effect size of 0.5). "G*Power 3.1" calculations for t-test tests indicated that a population size of 102 participants (= 2 groups x 51 participants in case equal participation per scenario) is needed. As shown in Table 7, for all of the scenarios there are more than 51 data points, and hence we conclude that t-test analysis results are statistically accurate for our study.[5]

t-test is one of the linear models (along with regression and ANOVA) used for statistical analysis and one of the assumptions of such models is the normality of the variables (Sainani, 2012). However, as Sainani (2012) contended that t-test is about making inferences about the means and hence the normality assumption is not critical due to the Central Limit Theorem if there is sufficient data and some scholars suggested (e.g., Le Cessie et al., 2020) t-test is appropriate if there are more than 25 observations and if there are no extreme outliers. As shown in Table 7 there are more than 25 observations for both review writers and review readers for all scenarios. We also checked the box plots of the review writers' and review readers' objective ambivalence values for all scenarios, presented in Figure 6. For scenario 1 (MTurk CA Readers - MTurk CA/US Writers) there is only 1 outlier out of 90 observations (for review writers – objective ambivalence is greater than the upper fence value of 6.5), for scenario 2 (MTurk US Readers - MTurk CA/US Writers) there are only 2 outliers out of 135 observations (for review writers – objective ambivalence is greater than the upper fence value of 7) , for scenario 3 (Student Readers - MTurk CA/US Writers)

[5] For all scenarios for an effect size of 0.5, at $\alpha = 0.05$ level calculated statistical powers (using G*Power 3.1 with the number of observations presented in Table 7) are more than 0.9.

there is only 1 outlier out of 228 observations (for review writers – objective ambivalence is greater than the upper fence value of 7) , and for scenario 4 (Student Readers - Student Writers) there are no outliers. We removed these review writers' responses and associated review readers' responses from the dataset. After removing the outliers, we conclude that using the t-test is appropriate for our study.

Since the study participants (i.e., review writers and review readers) form two separate groups we employed the independent t-test (Newman, 1980). We checked the variances of the variable of interest, i.e., objective ambivalence, for review writers and review readers for all of the 4 scenarios using the Levene's test and reported the t-test results with equal variances and stated otherwise if the variances are significantly different from each other.

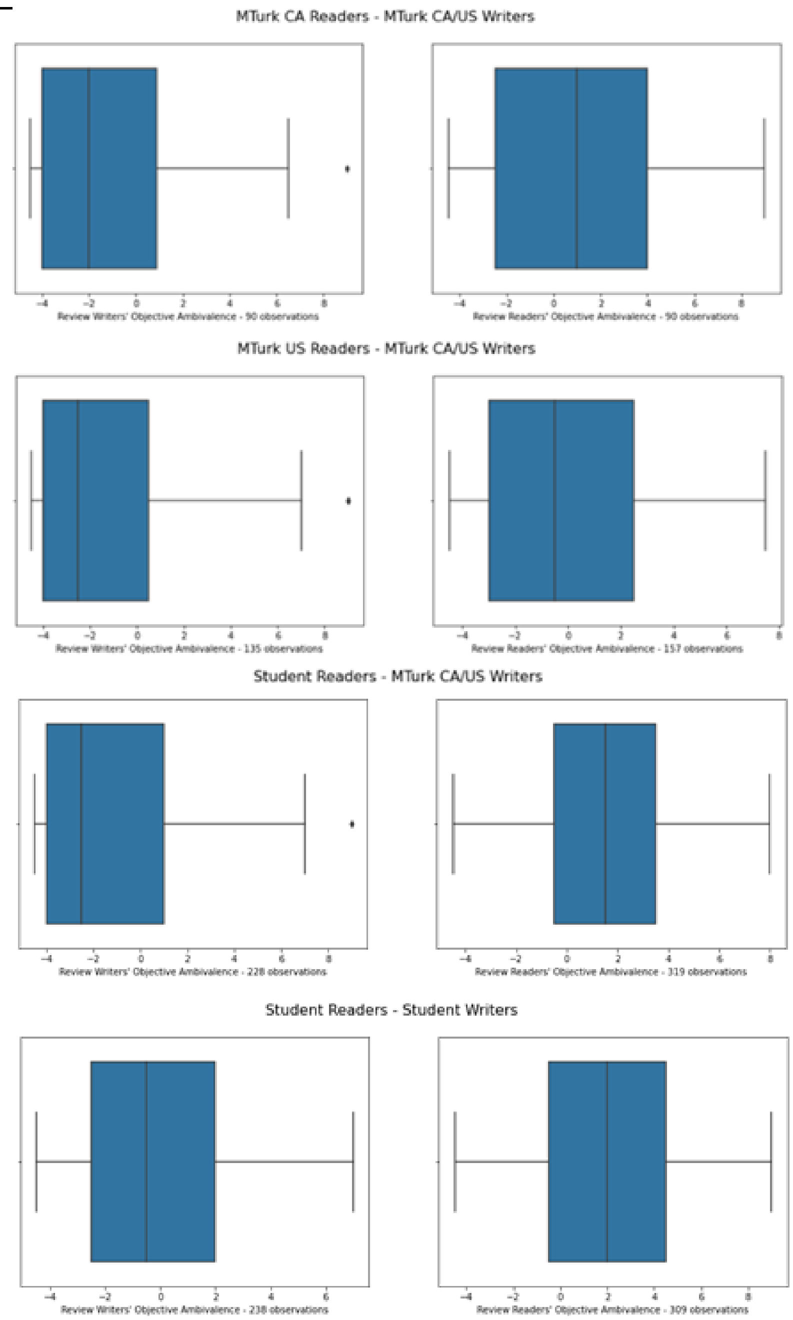

Figure 6 Box plots for review writers' (left) and review readers' (right) objective ambivalences for scenario 1 (top row), 2 (second row), 3 (third row) and 4 (bottom row)

4.2.1 Hypothesis Testing using Objective Ambivalence

Objective ambivalence is calculated using the SIM (Similarity-Intensity model) as suggested by Thompson et al. (1995):

$$Objective\ Ambivalence = \frac{PosEval\ +\ NegEval}{2} - abs\ (PosEval\ -\ NegEval)$$

where PosEval and NegEval represent the positive and negative evaluations of (towards) the restaurant in the review. They are measured using the following item "Considering only your POSITIVE (NEGATIVE) thoughts and feelings about this RESTAURANT and ignoring the negative (positive) ones, how positive (negative) would you say your positive thoughts and feelings are?" on a 10-point scale (See Appendix B for details).

The results of the t-tests, depicted in Table 10, support the stated hypothesis: review readers' ambivalent attitude towards the restaurant is higher than the review writers' ambivalent attitude towards the restaurant for all four scenarios.

Scenarios	obs # (reader)	obs # (writer)	mean reader obj ambivalence	mean writer obj ambivalence	difference (reader-writer)	t-value	p-value
MTurk CA Reader - MTurk CA/US Writer[6]	89	89	0.88	-1.12	2	4.19	**<0.01**
MTurk US Reader - MTurk CA/US Writer	155	133	-0.15	-1.44	1.29	3.29	**<0.01**
Student Reader - MTurk CA/US Writer	317	227	1.42	-1.24	2.66	9.82	**<0.01**
Student Reader - Student Writer	309	238	1.97	0.21	1.76	6.73	**<0.01**

Table 10 Summary of t-test hypothesis testing results using objective ambivalence.

[6] For this case the variance equality test was significant, i.e., the variances of review writers' and review readers' objective ambivalence evaluations are significantly different from each other. Hence, t-test is applied using un-equal variances.

4.2.2 Hypothesis Testing using Text Mining (Sentiment Analysis)

To further strengthen our findings, we also employed sentiment analysis to review writers' and review readers' texts using the VADER sentiment analysis tool. VADER provides multidimensional (i.e., positive, negative, and neutral) measures of a text. It also provides a compound score that can be used to identify the overall sentiment of the text. For this study, we only extracted the positive and negative scores for each review text using VADER. To test our hypothesis, we first used VADER to extract positive and negative emotions from the review texts (review text from review writers and review helpfulness text about the review from the review readers) and then conducted t-tests for all of the scenarios.

To measure the ambivalence from a given text, we employed the same SIM ambivalence model as stated above; specifically, we used the following formula to compute text ambivalence:

$$Text\ Ambivalence\ = \frac{pos +\ neg}{2} - |pos - neg|$$

where *pos* and *neg* are the positivity and negativity scores of a given text obtained using VADER. Before conducting the t-test we iteratively checked for the outliers in the data using the box plots presented in Figure 7. We checked and removed the outliers after removing the outliers for the self-reported values as explained in the previous section. As shown in the figure, there are some outliers in the data: for scenario 1 (MTurk CA Readers - MTurk CA/US Writers) there are 3 outliers out of 89 observations (for review readers – text objective ambivalence is less than the lower fence value of -0.1345 (2 data points) and after removing the outliers we noticed that there is still one more outlier: text objective ambivalence is less than the lower fence value of -0.1295 (1 data point)), for scenario 2 (MTurk US Readers - MTurk CA/US Writers) there are 2 outliers out of 155 observations (for review readers – text objective ambivalence is greater than the upper fence value of 0.107 and less than the lower fence value of -0.1625), for scenario 3 (Student Readers - MTurk CA/US Writers) there 2 outliers (1 for review readers – text objective ambivalence is lower than the lower fence value of -0.1545 and 1 for review writers – text objective ambivalence is lower than the lower fence value of -0.1945 [associated review readers responses are also removed]), and for scenario 4 (Student Readers - Student Writers) there are no outliers. We removed these review writers' responses and associated review readers' responses from the dataset. After removing the outliers, we conclude that using the t-test is appropriate for our study.

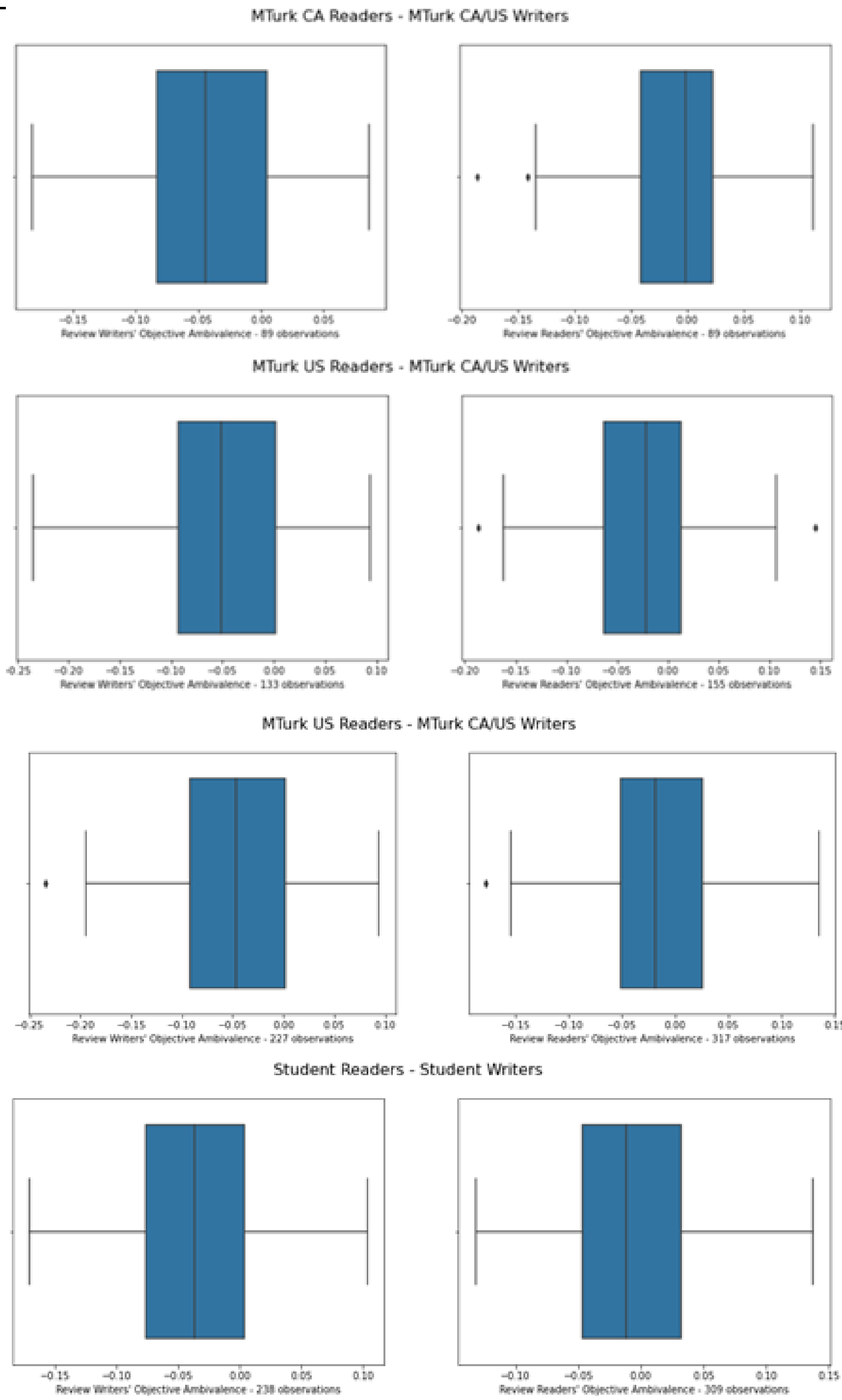

Figure 7 Box plots for review writers' (left) and review readers' (right) review text ambivalences (for writers) and review helpfulness text ambivalences (for readers scenario 1 (top row), 2 (second row), 3 (third row) and 4 (bottom row)

The results of our t-tests, depicted in Table 11, support our stated hypothesis: The anticipatory ambivalence of the receiver/reader based on a two-sided review content for a focal service is higher than the ambivalent attitude of the source/writer of the review who has already experienced the focal service.

Scenarios	obs # (reader)	obs # (writer)	mean reader obj ambivalence	mean writer obj ambivalence	difference (reader-writer)	t-value	p-value
MTurk CA Reader - MTurk CA/US Writer	86	86	-0.003	-0.044	0.041	4.627	**<0.01**
MTurk US Reader - MTurk CA/US Writer[7]	153	132	-0.025	-0.048	0.024	3.233	**<0.01**
Student Reader - MTurk CA/US Writer[7]	314	225	-0.014	-0.045	0.032	6.213	**<0.01**
Student Reader - Student Writer	309	238	-0.008	-0.035	0.027	5.626	**<0.01**

Table 11 Summary of t-test hypothesis testing results using text ambivalence.

[7] For this case the variance equality test was significant, i.e., the variances of review writers' and review readers' objective ambivalence evaluations are significantly different from each other. Hence, t-test is applied using un-equal variances.

5 Discussion

Online review platforms, such as Yelp and TripAdvisor, facilitate convenient access to a large number of online reviews (Qahri-Saremi & Montazemi, 2023). However, they impose an information load on the consumers due to their large number of reviews. To mitigate the information load, it is recommended that review websites provide more useful information to consumers (Wang et al., 2020) by providing more helpful reviews, as review helpfulness is a measure of perceived value for the consumer (Mudambi & Schuff, 2010).

Previous studies identified many factors affecting review helpfulness, including review star rating, review valence, review sidedness, and reviewer expertise. However, the findings are mixed for some factors, e.g., review star rating, review valence and review sidedness. Specifically for the review star rating, it is argued that it doesn't capture the reviewers' attitude accurately. For example, the star rating of a review, which is believed to measure a reviewer's attitude, cannot identify indifference and ambivalence: individuals who are indifferent or ambivalent are inclined to give a 3 star to a product/service (Klopfer & Madden, 1980). Also, within these studies, it is implicitly assumed that the attitude enacted by the reviewer via the review text is completely transferred to the review reader. Nonetheless, such an assumption may not be correct.

To the best of our knowledge, none of the extant studies investigated whether the review writers' attitudes are completely transferred to review readers. To ameliorate this void, in this study, we investigated the following research question:

Are the review writers' attitudes enacted in the review text completely transferred to the review readers?

In responding to the research questions, our research makes important contributions to theory and practitioners working in the field, as explained next.

5.1 Contributions to the Theory

The contribution of our findings to theory is significant. Our hypothesis that is supported by our empirical study contends that: The anticipatory ambivalence of the receiver/reader based on a two-sided review content for a focal service is higher than the ambivalent attitude of the source/writer of the review who has already experienced the focal service. As a result, the findings from all the extant studies that used review writers' star ratings as a proxy for the review readers' emotions/attitudes are questionable. We delve into the details of this statement as follows.

5.1.1 Review Star Rating and Review Helpfulness

Although consumers can benefit from a vast number of reviews available on consumer review sites, the amount of information makes it difficult for the consumer to process and judge reviews (Malhotra, 1984). Since consumers have limited cognitive and resource capacity resources (Roetzel, 2019), reviewing websites should provide more useful information to consumers (Wang et al., 2020). To that end, there has been extensive research in identifying factors that make a review helpful (Qahri-Saremi & Montazemi, 2019), and one such factor is the review star rating. Star rating shows the reviewer's overall evaluation of a product/service (Lopes et al. 2020), where 1 star represents an extremely negative experience, and 5 stars represent an extremely positive experience of the product/service (Mudambi & Schuff, 2010). Although the review star rating is one of the most used variables in the assessment of review helpfulness, the findings about the relationship between the review star rating and review helpfulness are mixed (See Table 12). The implicit assumption in all these studies is that after reading the review, the emotions of the review writer are completely transferred to the reader. Therefore, they expect a relationship between the review writer's overall emotion represented by his/her star rating of the product/service and the reader's overall emotion gleaned from reading the review and represented by his/her usefulness rating of the review. We contend that this assumption is flawed. Our analyses in support of our hypothesis show that the sender and receiver of the review have two different prospect-based emotions. The prospect-based emotions are "characterized as reactions to (i.e., having a positive or negative feeling about) an envisaged [anticipatory] event [i.e., anticipatory ambivalence by the reader of the review], or to the confirmation/disconfirmation of the prospect of such an event" (Ortony et al., 2022, p.126), (i.e., ambivalent attitude) by the writer/sender of the review. Therefore, a reason for mixed findings of the relationship between star rating and helpfulness is disregard for the two different prospect-based emotions of the review sender and receiver. To test this conjecture, we asked our review writer subjects to provide their overall impression of the focal restaurant using a scale 1 to 5 (where 1 star represents an extremely negative experience, and 5 stars represent an extremely positive experience. See APPENDIX – A Survey Form for Online Review Writers). Furthermore, the review readers' subjects were asked to provide their overall impression of the focal restaurant based solely on the review that they received, using a scale of 1 to 5. It should be noted that the review readers did not have access to the star rating of the review writers. Our t-test analyses (depicted in Table 13)[8] show that the star rating of the review readers is significantly

---

[8] For this case the variance equality test was significant, i.e., the variances of review writers' and review readers' star ratings are significantly different from each other. Hence, t-test is applied using un-equal variances.

lower than review writers. Before conducting the t-test we checked the whole data for outliers using boxplots, depicted in Figure 8, and removed 30 data points from the review writers' star rating assignment (and associated review readers' review star rating assignment) and removed 59 (in total) data points from the review readers' review star rating assignment due to outliers.

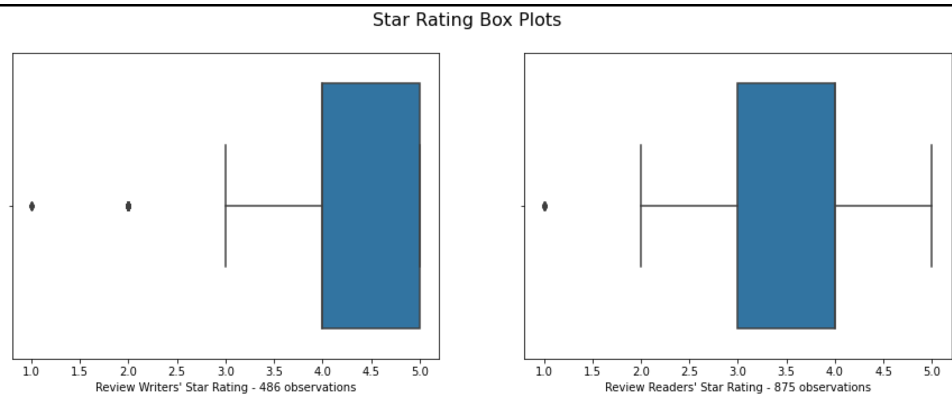

Figure 8 Box plots for review writers' (left) and review readers' (right) star rating assessments

This finding supports our conjecture, that implicit assumption that after reading the review, the emotions of the review writer are completely transferred to the reader. Therefore, it is quite likely that mixed findings for the relationship between star rating provided by the review writer and review reader impression of the review helpfulness is at least partly due to the flawed assumption in the extant studies depicted in Table 12.

Authors	Findings	Goods/Service Type
Zhang, Craciun et al. (2010)	mixed(positive/negative)	experience (photo-editing software and anti-virus software)
Sousa & Pardo (2021)	mixed(positive/negative) (correlation analysis)	experience
Xu et al. (2023)	negative	not specified(reviews from Yelp)
Kwok & Xie (2016)	negative	experience

Jayasingh & Thiagarajan (2022)	negative	search (mobile phone)
Yin, Bond et al. (2014)	negative	experience
Racherla & Friske (2012)	negative/no relationship	experience, search and credence
Zhou & Guo (2015)	negative	experience
Sun et al. (2019)	negative	experience and search
Zhou & Yang (2019)	negative	experience and search
Fan & Zhang (2020)	negative	search (mobile phone)
Li, Liu et al. (2020)	negative	experience
Shaft et al. (2020)	negative	experience
Mariani & Borghi (2021)	negative	experience
Li et al. (2021)	negative	experience
Liang et al. (2019)	negative	experience
Zhu et al. (2014)	negative	experience
Guo & Zhou (2017)	negative	experience
Zhou & Guo (2017)	negative	experience
Li, Wang, et al. (2019)	negative	experience
Wang, Tang et al. (2019)	negative	experience
Hu et al. (2017)	negative	experience
Chua & Banerjee (2015)	negative	experience (books)
Yin, Zhang et al. (2014)	negative	experience
Wang & Karimi (2019)	negative	experience and search (printer, tv, book, music album) (no separate analysis)
Wang,Li et al. (2015)	negative	experience and search (no separate analysis)
Karimi & Wang (2017)	positive	experience (mobile games)
Pan & Zhang (2011)	positive	experience and search (e.g., music CDs, movie DVDs, video games, GPS, photo-editing software, food supplements) (no separate analysis)
Liu & Park (2015)	positive	experience
Yin et al. (2017)	mixed(positive/negative)	experience (Apple’s App Store Apps)
Chatterjee (2020)	positive	experience
Biswas et al. (2021)	positive	experience and search (no separate analysis)

Xu et al. (2022)	positive	experience (movies)
Korfiatis et al. (2012)	positive	experience (books)
Huang et al. (2015)	positive	experience and search (a cell phone, printer, camera, music player, music CD, and video game) (no separate analysis)
Ullah et al. (2015)	positive	experience (movies)
Quaschning et al. (2015)	positive	experience (book)
Bjering et al. (2015)	positive	experience and search (no separate analysis)
Ahmad & Laroche (2015)	positive	experience(kitchen appliances)
Agnihotri & Bhattacharya (2016)	positive	experience and search (no separate analysis)
Ren & Hong (2019)	positive	experience and search
Yang et al. (2020)	positive	experience (reviews from amazon: beauty, grocery, cell phone, clothing, and video products)
Craciun et al. (2020)	positive	search (tablet computer)
Zhao (2020)	positive	not specified(reviews from Amazon)
Willemsen et al. (2011)	positive	experience and search (no separate analysis)
Wang, Wang et al. (2019)	positive	meta analysis
Hong et al. (2017)	positive/no relationship	meta analysis (experience and search)

Table 12 Studies investigated the relationship between review star rating and review helpfulness.

	review reader	review writer	mean diff	t-test	p-value
number of observations	816	456			
Review star rating	3.86	4.15	-0.29	-6.66	0<0.01

Table 13 Mean review star rating evaluations of review readers and review writers.

5.1.2 Gender Differences in the Knowledge/Emotion Transfer

Extant literature shows that emotion norms are different between men and women, and generally, women express more emotions than men (Craciun & Moore, 2019). Hence, to check if there are any differences in the evaluations of men and women we did a t-test analysis for the objective ambivalence, review star, positive and negative restaurant evaluation variables. For the analysis we used the data where participants (review readers) provided their gender, and we removed any duplicates from the data. Since review readers can read online reviews written by individuals with various demographic characteristics and backgrounds, it is the review readers' assessment that creates the difference between the review writers' and review readers' assessment. Hence, we conducted our analysis grouped by the review readers' gender irrespective of the review writers' gender.

Table 14 provides the t-test results for the objective ambivalence values for review readers grouped by gender. Prior to t-test we utilized box plots, depicted in Figure 9, to remove outliers in the whole dataset. Based on the box plots we removed 2 data points from the review writers which has more than 7 (upper fence value) objective ambivalence. We also removed associated review readers' responses (2 responses from female review readers and 3 responses from male review readers).

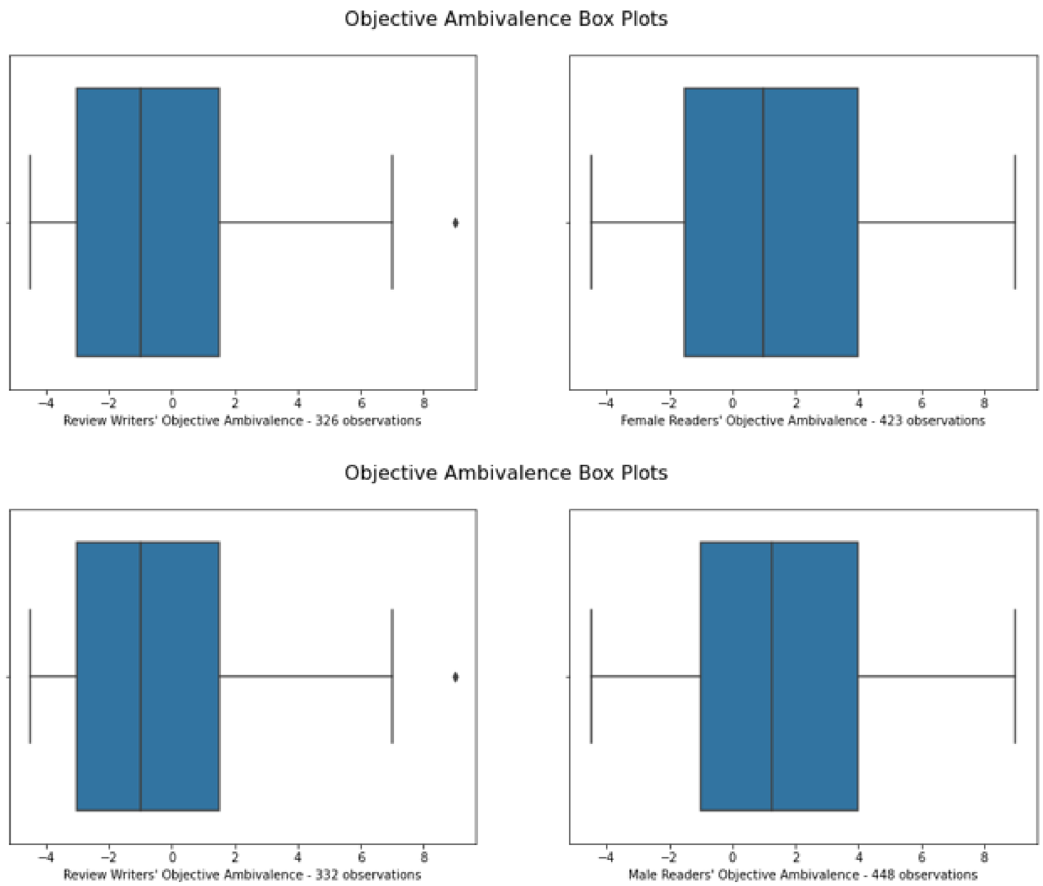

Figure 9 Box plots for review writers' (left) and review readers' (right) objective evaluations for female readers (top) and for male readers (bottom)

The results support our hypothesis: as the results indicate gender doesn't play a role in the transfer of attitudes as review readers (both men and women) feel more ambivalent than review writers (irrespective of gender) who used the restaurant service before.

	obs # (reader)	obs # (writer)	mean reader obj ambivalence	mean writer obj ambivalence	difference (reader-writer)	t-value	p-value
Female Reader	421	325	1.18	-0.60	1.78	7.34	**<0.01**
Male Reader	445	331	1.38	-0.71	2.09	9.18	**<0.01**

Table 14 Mean objective ambivalence of review readers and review writers grouped by gender.

Furthermore, review readers also devalue the review star rating assigned by the review writers. In line with the previous analysis, we plotted boxplots, depicted in Figure 10 , to remove outliers in the whole dataset. Based on the box plots we removed data points from the review writers which have less than 3 (lower fence value) star rating and from review readers which have less than 2 (lower fence value) star rating. In total for the female reader group, we removed 28 responses from the readers and 22 responses from the review writers, and for the male reader group we removed 31 responses from the readers and 20 responses from the review writers.

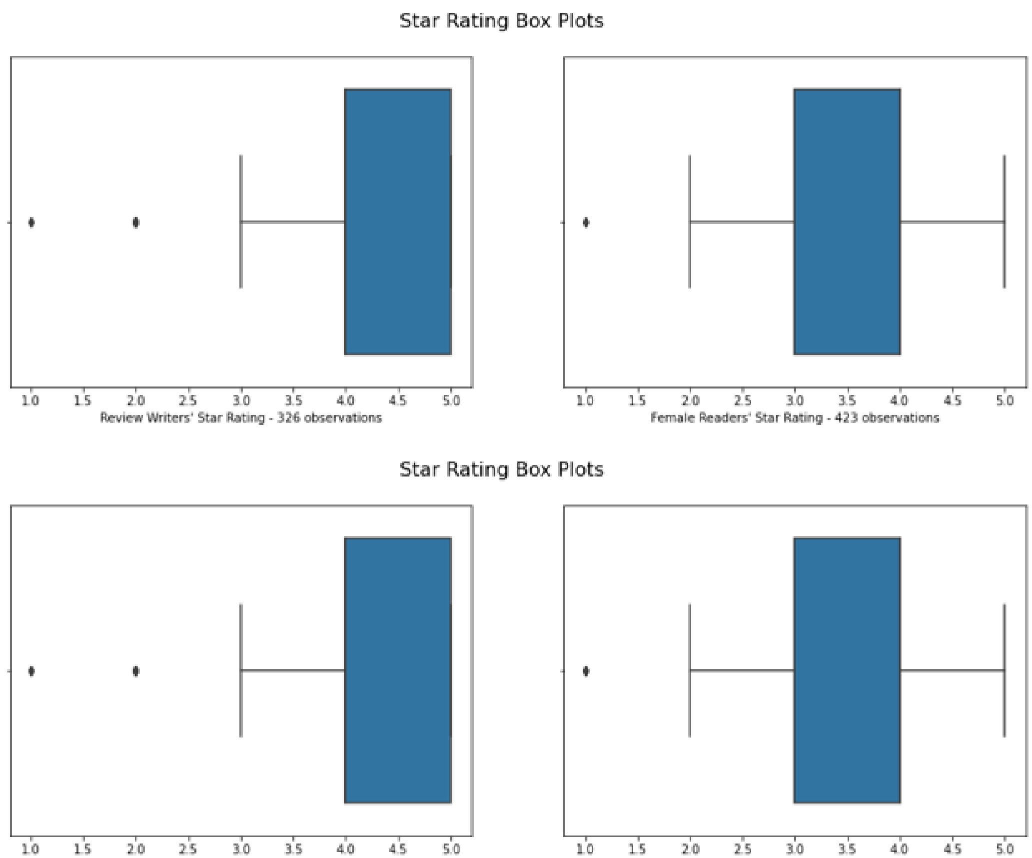

Figure 10 Box plots for review writers' (left) and review readers' (right) star rating assignments for female readers (top) and for male readers (bottom)

Table 15 presents the review star ratings for review readers and review writers grouped by gender. As the results indicate, the star rating of the review readers is significantly lower than review writers irrespective of gender.

	obs # (reader)	obs # (writer)	mean reader star rating	mean writer star rating	difference (reader-writer)	t-value	p-value
Female Reader[9]	395	304	3.86	4.17	-0.31	-5.54	**<0.01**
Male Reader	17	312	3.87	4.18	-0.31	-5.54	**<0.01**

Table 15 Mean review star rating evaluations of review readers and review writers grouped by gender.

We also checked if gender affects positive and negative evaluations using t-tests. Prior to the t-test we utilize box plots, depicted in Figure 11 (for female readers) and Figure 12 (for male readers), to remove outliers in the data. We iteratively used boxplots to ensure that there are no outliers in the whole dataset, and we checked positive and negative evaluations independently and we removed a data point if it is an outlier either for positive or negative evaluation. Based on the boxplots for the female reader group we retained data where the review writers' positive evaluation is at least 6 (out of 9), review readers' positive evaluation is at least 3 (out of 9) and review writers' negative evaluation is at most 5 (out of 9) (based on iterative outlier removal). For the male reader group, we retained data where the review writers' positive evaluation is at least 6 (out of 9), review readers' positive evaluation is at least 3 (out of 9) and review writers' negative evaluation is at most 7 (out of 9) (based on iterative outlier removal). In total for the female reader group, we removed

[9] For female readers the variance equality test was significant, i.e., the variances of review writers' and review readers' star ratings are significantly different from each other. Hence, t-test is applied using unequal variances.

86 data points from the review readers and removed 66 data points from the review writers, and for the male reader group we removed 79 data points from the review readers and removed 52 data points from the review writers.

After the data cleaning, to check if gender affects the evaluations we did the t-test by grouping the review readers based on the gender. The analysis results, provided in Table 16[10], indicate that regardless of gender, review readers are more conservative, i.e., more negative and less positive, compared to the review writers. Considering all of the analysis related with the gender, we contend that our analysis is robust, and regardless of review reader gender, our hypothesis that review readers are more ambivalent than review writers hold.

[10] For these cases the variance equality test was significant, i.e., the variances of review writers' and review readers' positive/negative restaurant evaluations are significantly different from each other. Hence, t-test is applied using un-equal variances.

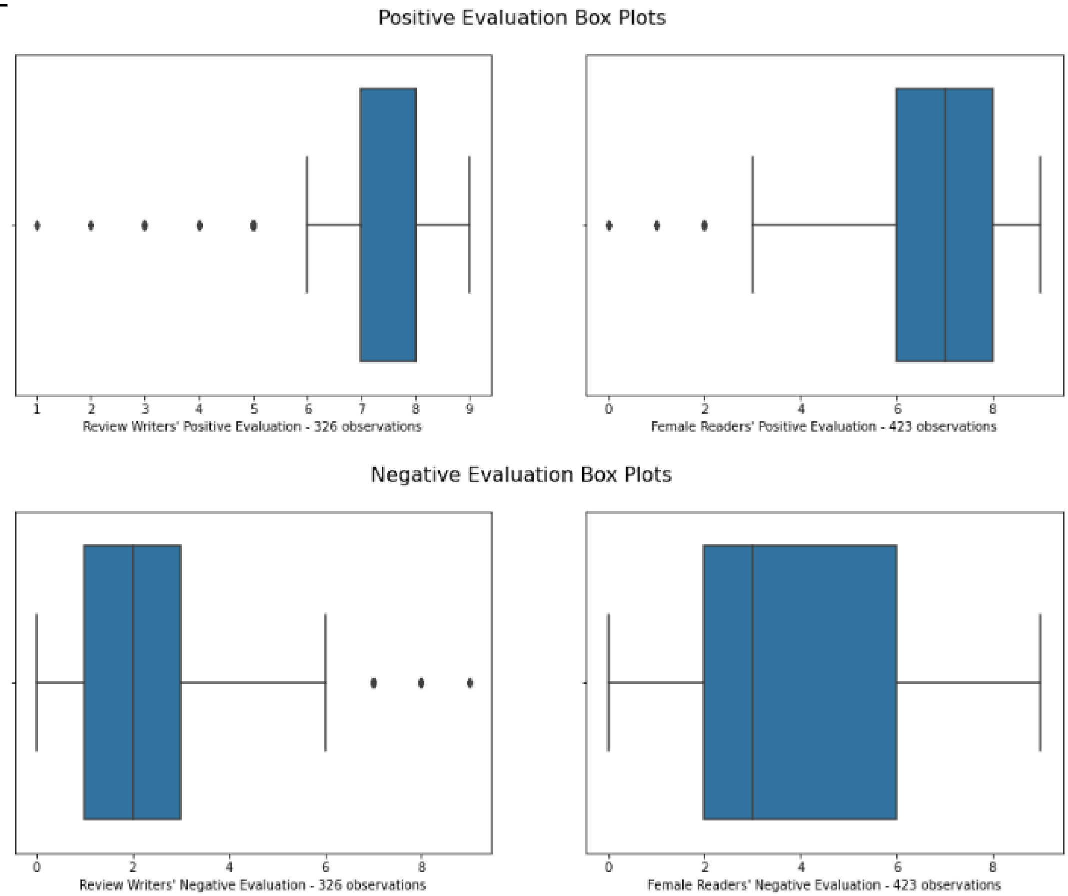

Figure 11 Box plots for review writers' (left) and female review readers' (right) positive evaluations (top) and negative evaluations (bottom)

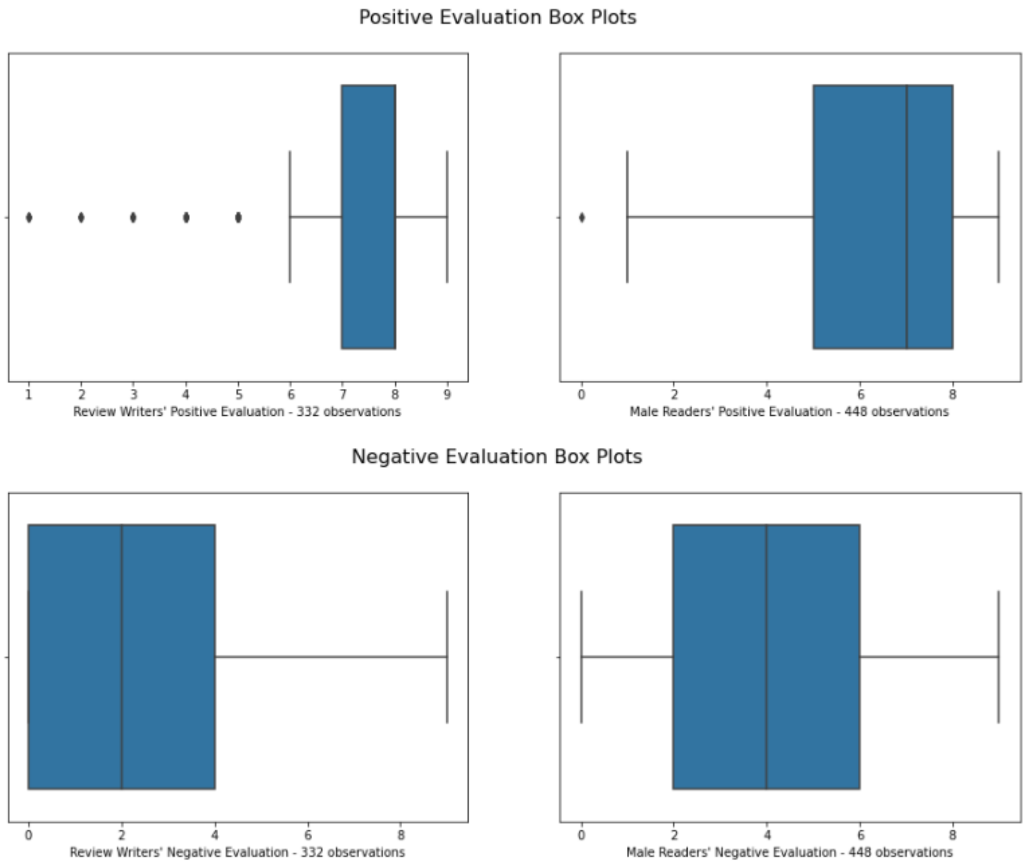

Figure 12 Box plots for review writers' (left) and male review readers' (right) positive evaluations (top) and negative evaluations (bottom)

	obs # (reader)	obs # (writer)	mean reader positive eval.	Mean writer positive eval.	Difference (reader-writer)	t-value	p-value
Female Reader	337	260	6.99	7.75	-0.76	-7.54	**<0.01**
Male Reader	369	280	6.75	7.70	-0.95	--9.89	**<0.01**
	obs # (reader)	obs # (writer)	mean reader negative eval.	mean writer negative eval.	difference (reader-writer)	t-value	p-value
Female Reader	337	260	3.31	1.45	1.86	11.87	**<0.01**
Male Reader	369	280	3.57	1.73	1.84	11.49	**<0.01**

Table 16 Mean positive/negative evaluations of review readers grouped by gender.

5.2 Implications for Practice

Our study provides valuable implications for practice. Firstly, our findings show that there is a significant difference in the review readers' and review writers' evaluations of the restaurants stated in the reviews. We conducted t-tests to assess the differences between the review writers' and review readers' positive and negative evaluations of the restaurants. Before conducting the t-tests we utilized the boxplots, depicted in Figure 13, for outlier removal from the whole dataset. We iteratively used boxplots to ensure that there are no outliers in the data, and we checked positive and negative evaluations independently and we removed a data point if it is an outlier either for positive or negative evaluation. Based on the boxplots we retained data where the review writers' positive evaluation is at least 6 (out of 9), review readers' positive evaluation is at least 3 (out of 9) and review writers' negative evaluation is at most 7 (out of 9) (based on iterative outlier removal). Based on the boxplots we removed 148 data points from the review readers and removed 75 data points from the review writers.

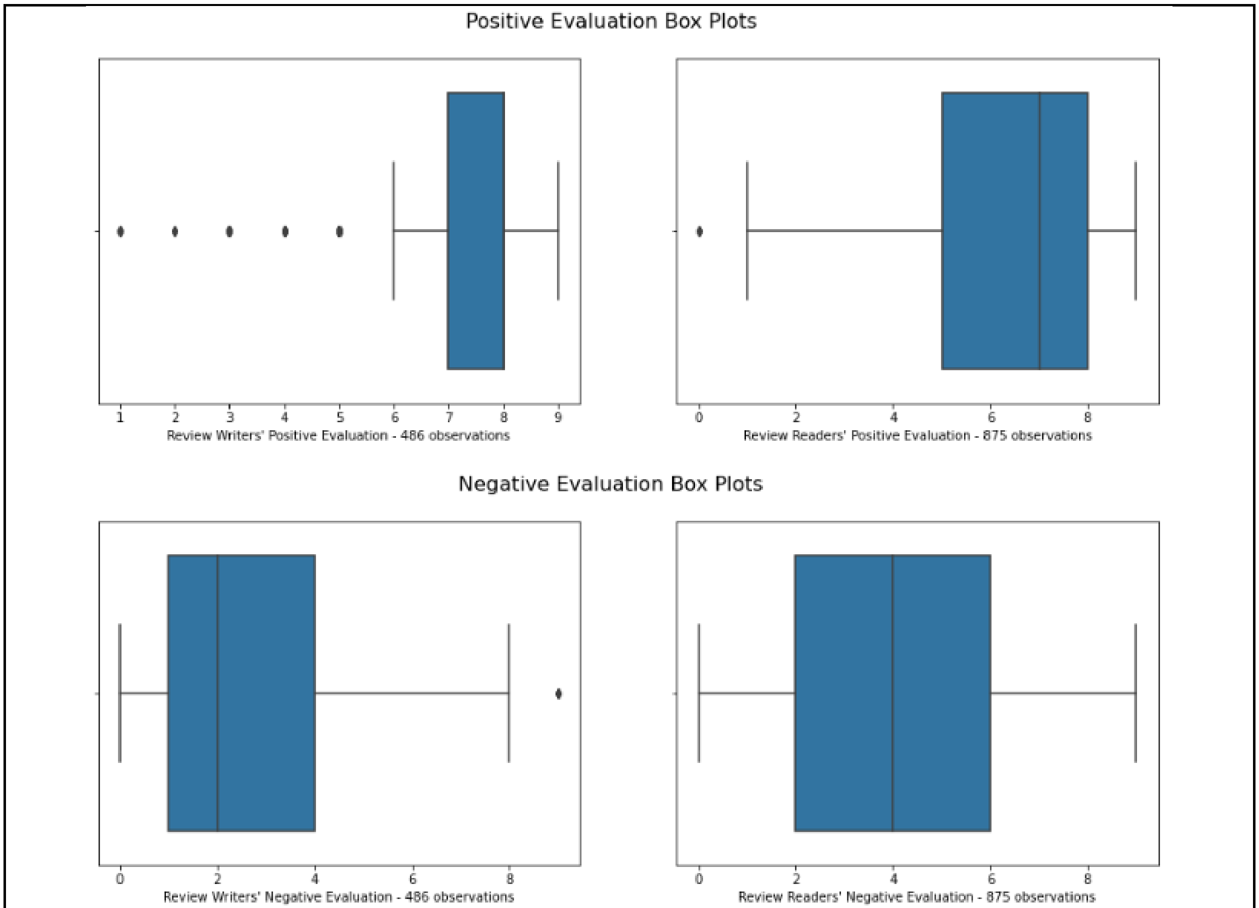

Figure 13 Box plots for review writers' (left) and review readers' (right) positive evaluations (top) and negative evaluations (bottom)

As shown in Table 17[11], review readers significantly overvalue negativity and devalue the positivity of the review writers' evaluations. That is, while review writers' mean negative (positive)

[11] For this case the variance equality test was significant, i.e., the variances of review writers' and review readers' positive/negative restaurant evaluations are significantly different from each other. Hence, t-test is applied using un-equal variances.

evaluation of the restaurant is 1.80 (7.71), review readers' mean negative (positive) evaluation of the same restaurants is 3.46 (6.85): Review readers' negative (positive) evaluation are significantly higher (lower) than the review writers. Extant research finds that, compared to positive reviews, negative reviews have significantly more effect in adversely affecting the consumer attitudes, purchase intentions and sales which is a consequence of a decrease in the perceived reliability of the seller (Septianto et al., 2020). Consequently, service providers should pay more attention to negative messages in the reviews and need to mitigate factors resulting in negative reviews to reverse the adverse effect of negativity on their services. This is even more urgent than previously assumed since star ratings on the websites (e.g., Yelp.com) reflect the attitude of the review writer who has already used/purchased the service. Our findings show that the review reader, who is at the pre-purchase state, experiences more negative emotion/attitude than review writer. Thus, a star rating of 4 that was assumed to be good may not hold because the review reader attitude after reading the review could be significantly lower (e.g., star rating of 2) for the focal service.

	review reader	review writer	mean diff	t-test	p-value
Number of observations	727	411			
Negative evaluation of the restaurants	3.45	1.80	1.66	13.40	0<0.01
Positive evaluation of the restaurants	6.85	7.71	-0.85	-11.82	0<0.01

Table 17 Review readers and writers mean negative and positive evaluations of the restaurants.

Secondly, review websites should provide more helpful information to consumers (Wang et al., 2020) to decrease the information load (Roetzel, 2019). Several studies used review star rating as either an independent variable or a control variable in assessing the relationship between review star rating and review helpfulness; however, the findings are inconclusive (See Table 12). We contended that this is due to the star rating's inability to accurately capture the reviewers' underlying positive and negative evaluations (Resnick et al., 2000). Hence, we propose review websites collect ambivalence attitude (e.g., questionnaire similar to the one used in this research depicted in APPENDIX – A Survey Form for Online Review Writers) instead of review star to capture review writers' underlying sentiment about the focal product/service. Furthermore, they should capture the ambivalent attitude of the review reader by means of questionnaires (see APPENDIX – B Survey Form for Online Review Readers) to better understand their emotional state towards purchase of the focal product/service.

5.3 Limitations & Future Research

Firstly, this study examined the experience services, i.e., restaurants, in the assessments of attitude/emotion transfer from the review writers to review readers. Experience services, compared to search services, can only be subjectively evaluated after consuming the service (Xiao & Benbasat, 2007) creating uncertainty in the service quality assessment (Wang et al., 2021) and hence increasing the consumers' perceived risk (Zhang, Wang, Wu, 2021). However, search services can be objectively evaluated before using them (Xiao & Benbasat, 2007). Examples of search services include checking account or credit card for banks (Licata & Frankwick, 1996) or selecting a cable TV operator (Lima & Fernandes, 2015). The question in need of future empirical research is whether the ambivalent attitude enacted by the reviewer in the review text is completely transferred to review readers or not?

Secondly, we conducted our empirical investigation within the Canadian and US context. As Hong et al. (2016) contend, culture is one of the factors that affect consumer behaviour, and compared to consumers from collectivist cultures (e.g., China), consumers from individualist cultures (e.g., USA and Canada) are more likely to include more emotional expressions in their reviews (Hong et al., 2016). Future research can assess whether our findings hold in collectivist cultures.

APPENDICES

APPENDIX – A Survey Form for Online Review Writers

Summation Check X + Y= ? User Input

1. What is your age? :
2. MTurk/Student ID
3. What is your gender?

Gender	Response	
Female	☐	
Male	☐	
Prefer to specify	☐	
Prefer not to answer	☐	

4. What is your highest completed level of education?

Education Level	Response
Below High School	☐
Some high school education	☐
High school Diploma	☐
Undergraduate (bachelor's) Degree	☐
Master's Degree	☐
Doctoral Degree	☐
Prefer not to answer	☐

5. Please indicate how much the following statements apply to you. *

		1	2	3	4	5	6	7	
My thoughts are often contradictory	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
Many topics make me feel conflicted	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
I usually see both the positive as well as the negative side of things	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
I often experience both sides of an issue pulling on me	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
I often find that there are pros and cons to everything	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
I often feel torn between two sides of an issue	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
Most of the time, my thoughts and feelings are not necessary in accordance with each other	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided
Sometimes when I think about a topic, it almost feels like I am physically switching from side to side	Not At All Undecided	○	○	○	○	○	○	○	Very Undecided

* Data from this question is not used in this research.

My feelings are often simultaneously positive and negative	Not At All Undecided 1 2 3 4 5 6 7 Very Undecided
I often experience that my thoughts and feelings are in conflict when I'm thinking about a topic	Not At All Undecided 1 2 3 4 5 6 7 Very Undecided

6. Consider a restaurant that you visited, and your experience was [One of the following three choices will be displayed in random]:
   i- Overall positive
   ii- Overall negative
   iii- Mixed (Mix of positive and negative)

   Please write a review about your experience with that restaurant. **The review should contain at least 100 words**.

a. Name of the Restaurant (as shown on TripAdvisor):

b. Province/State of the Restaurant
c. City of the Restaurant

d. Price Range of the Restaurant (as shown on TripAdvisor)

e. Approximate visit date

f. Please write a positive and negative review about your experience with that restaurant. **The review should contain at least 100 words**:

Please write your review here.

Please indicate your evaluations of the **restaurant visit:**

i- Considering only your **POSITIVE** thoughts and feelings **about your restaurant visit** and ignoring the negative ones, how positive would you say your positive thoughts and feelings are?

(0: No positive thoughts or feelings, 9: Maximum positive thoughts or feelings)

0	1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○	○

ii- Considering only your **NEGATIVE** thoughts and feelings **about your restaurant** visit and ignoring the positive ones, how negative would you say your negative thoughts and feelings are?

(0:No negative thoughts or feelings, 9:Maximum negative thoughts or feelings)

0	1	2	3	4	5	6	7	8	9
○	○	○	○	○	○	○	○	○	○

g. Please indicate your overall rating of the restaurant:[*]

	1	2	3	4	5	
Extremely Negative	○	○	○	○	○	Extremely Positive

h. To what extent do you **feel undecided** about how good or bad **your restaurant visit** was?*

	1	2	3	4	5	6	7	
Not At All Undecided	○	○	○	○	○	○	○	Very Undecided

i. To what extent do you **feel conflicted** about how good or bad **your restaurant visit** was?*

	1	2	3	4	5	6	7	
Not At All Conflicted	○	○	○	○	○	○	○	Very Conflicted

[*] Data from this question is not used in this research.

j. To what extent do you have **mixed feelings** about how good or bad **your restaurant visit** was?

	1	2	3	4	5	6	7	
Not At All Mixed	○	○	○	○	○	○	○	Very Mixed

7. Please indicate how much the following statements apply to you. *

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I see myself as creative	○	○	○	○	○	○	○
I see myself as imaginative	○	○	○	○	○	○	○
I see myself as unconventional	○	○	○	○	○	○	○
I see myself as moody	○	○	○	○	○	○	○
I see myself as easily upset	○	○	○	○	○	○	○
I see myself as anxious	○	○	○	○	○	○	○
I see myself as sympathetic	○	○	○	○	○	○	○
I see myself as warm	○	○	○	○	○	○	○
I see myself as kind	○	○	○	○	○	○	○
I see myself as dependable	○	○	○	○	○	○	○
I see myself as self-disciplined	○	○	○	○	○	○	○
I see myself as organised	○	○	○	○	○	○	○
I see myself as extraverted	○	○	○	○	○	○	○
I see myself as enthusiastic	○	○	○	○	○	○	○
I see myself as talkative	○	○	○	○	○	○	○

8. Please indicate if you have written a review for this restaurant before.

Have you written a review for this restaurant before?

Yes	No
○	○

* Data from this question is not used in this research.

9. Please indicate how often you post online reviews about your restaurant experiences.

How often do you write a review?

1 Never	2	3	4	5	6	7 Very Frequently
○	○	○	○	○	○	○

10. Please indicate the platforms you use review writing for restaurants.

Which platform do you use review writing and reading?

- ○ Tripadvisor
- ○ Yelp
- ○ Amazon
- ○ Facebook
- ○ Twitter
- ○ Google
- ○ Other

APPENDIX – B Survey Form for Online Review Readers

Summation Check X + Y= ? User Input

1. What is your age?:
2. MTurk/ Student ID
3. What is your gender?

Gender	Response	
Female	☐	
Male	☐	
Prefer to specify	☐	
Prefer not to answer	☐	

4. What is your highest completed level of education?

Education Level	Response
Below High School	☐
Some high school education	☐
High school Diploma	☐
Undergraduate (bachelor's) Degree	☐
Master's Degree	☐
Doctoral Degree	☐
Prefer not to answer	☐

5. Please indicate how much the following statements apply to you.*

Statement		1	2	3	4	5	6	7	
My thoughts are often contradictory	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
Many topics make me feel conflicted	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
I usually see both the positive as well as the negative side of things	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
I often experience both sides of an issue pulling on me	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
I often find that there are pros and cons to everything	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
I often feel torn between two sides of an issue	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
Most of the time, my thoughts and feelings are not necessary in accordance with each other	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided
Sometimes when I think about a topic, it almost feels like I am physically switching from side to side	Not At All Undecided	O	O	O	O	O	O	O	Very Undecided

* Data from this question is not used in this research.

My feelings are often simultaneously positive and negative	Not At All Undecided 1 2 3 4 5 6 7 Very Undecided
I often experience that my thoughts and feelings are in conflict when I'm thinking about a topic	Not At All Undecided 1 2 3 4 5 6 7 Very Undecided

6. The following review is written by a very knowledgeable reviewer for a mid-ranged priced restaurant:

Review:

Review is displayed here.

Using the scales below, how would you describe the above consumer review?

Using the scales below, how would you describe the above consumer review?

not at all helpful 1	2	3	4	5	6	7 very helpful
not at all useful 1	2	3	4	5	6	very useful 7
not at all informative 1	2	3	4	5	6	very informative 7

7. Please indicate your evaluations of **the review:**

i- Considering only your **POSITIVE** thoughts and feelings **about the review** and ignoring the negative ones, how positive would you say your positive thoughts and feelings are?

(0: No positive thoughts or feelings, 9: Maximum positive thoughts or feelings)

0 1 2 3 4 5 6 7 8 9

ii- Considering only your **NEGATIVE** thoughts and feelings **about the review** visit and ignoring the positive ones, how negative would you say your negative thoughts and feelings are?

(0:No negative thoughts or feelings, 9:Maximum negative thoughts or feelings)

0 1 2 3 4 5 6 7 8 9

8. To what extent do you **feel undecided** about how good or bad **this review** is? [*]

1 2 3 4 5 6 7

Not At All Undecided — Very Undecided

To what extent do you **feel conflicted** about how good or bad **this review** is?

1 2 3 4 5 6 7

Not At All Conflicted — Very Conflicted

To what extent do you **have mixed feelings** about how good or bad **this review** is?

1 2 3 4 5 6 7

Not At All Mixed — Very Mixed

[*] Data from this question is not used in this research.

9. Please explain the logic behind your evaluation of review helpfulness: (Please write at **least 100 words: positive and negative separately**):

Please write your explanation here.

10. Please indicate how much the following statements apply to you. *

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I see myself as creative	○	○	○	○	○	○	○
I see myself as imaginative	○	○	○	○	○	○	○
I see myself as unconventional	○	○	○	○	○	○	○
I see myself as moody	○	○	○	○	○	○	○
I see myself as easily upset	○	○	○	○	○	○	○
I see myself as anxious	○	○	○	○	○	○	○
I see myself as sympathetic	○	○	○	○	○	○	○
I see myself as warm	○	○	○	○	○	○	○
I see myself as kind	○	○	○	○	○	○	○
I see myself as dependable	○	○	○	○	○	○	○
I see myself as self-disciplined	○	○	○	○	○	○	○
I see myself as organised	○	○	○	○	○	○	○
I see myself as extraverted	○	○	○	○	○	○	○
I see myself as enthusiastic	○	○	○	○	○	○	○
I see myself as talkative	○	○	○	○	○	○	○

* Data from this question is not used in this research.

11. Please indicate the extent to which you have used online reviews to select a restaurant

When deciding on which restaurant to go to, I read online reviews

1 Never	2	3	4	5	6	7 Always
○	○	○	○	○	○	○

12. Please indicate the platforms that you have used in the past to read restaurant.

Which platform do you use review writing and reading?

- ○ Tripadvisor
- ○ Yelp
- ○ Amazon
- ○ Facebook
- ○ Twitter
- ○ Google
- ○ Other

APPENDIX – C Descriptive statistics

			Review Writer		Review Reader	
MTurk CA Reader - MTurk CA/US Writer			Sample Number	Sample Percentage	Sample Number	Sample Percentage
			90	100%	90	100%
	Gender					
		Female	42	46.67%	43	47.78%
		Male	47	52.22%	47	52.22%
		Other	1	1.11%	0	0.00%
		Prefer not to answer	0	0.00%	0	0.00%
		Total	90	100.00%	90	100.00%
			Mean	Range	Mean	Range
	Age		35.39	20-64	34.1	18-69
MTurk US Reader - MTurk CA/US Writer			**Review Writer**		**Review Reader**	
			Sample Number	Sample Percentage	Sample Number	Sample Percentage
			135	100%	157	100%
	Gender					
		Female	68	50.37%	68	43.31%
		Male	65	48.15%	88	56.05%
		Other	1	0.74%	1	0.64%
		Prefer not to answer	1	0.74%	0	0.00%
		Total	135	100.00%	157	100.00%
			Mean	Range	Mean	Range
	Age		35.47	20-66	39.54	21-69

Table 18 Descriptive statistics for MTurk CA/US Readers and CA/US Writers

Student Reader - MTurk CA/US Writer

		Review Writer		Review Reader	
		Sample Number	Sample Percentage	Sample Number	Sample Percentage
		228	100%	319	100%
Gender					
	Female	113	49.56%	156	48.90%
	Male	113	49.56%	161	50.47%
	Other	0	0.00%	0	0.00%
	Prefer not to answer	2	0.88%	2	0.63%
	Total	228	100.00%	319	100.00%
		Mean	Range	Mean	Range
Age		35.14	19-66	19.36	18-24

Table 19 Descriptive statistics for Student Writers and Student Readers (1 review reader hasn't specified an option for gender; hence it is counted as prefer not to answer)

Student Reader - Student Writer

		Review Writer		Review Reader	
		Sample Number	Sample Percentage	Sample Number	Sample Percentage
		238	100%	309	100%
Gender					
	Female	110	46.22%	156	50.49%
	Male	127	53.36%	152	49.19%
	Other	0	0.00%	0	0.00%
	Prefer not to answer	1	0.42%	1	0.32%
	Total	238	100.00%	309	100.00%
		Mean	Range	Mean	Range
Age		19.29	18-30	19.62	18-25

Table 20 Descriptive statistics for Student Writers and Student Readers

APPENDIX – D McMaster University Research Ethics Board Approval

McMaster University | Research Ethics Board

McMaster University Research Ethics Board (MREB)
c/o Research Office for Administrative Development and Support
MREB Secretariat, GH-305
1280 Main St. W.
Hamilton, Ontario, L8W 4L8
email: ethicsoffice@mcmaster.ca
Phone: 905-525-9140 ext. 23142

CERTIFICATE OF ETHICS CLEARANCE TO INVOLVE HUMAN PARTICIPANTS IN RESEARCH

Today's Date: Jul/16/2021

Supervisor: Dr. Ali Reza Montazemi
Student Investigator: Mr. Mehmet Akgul
Applicant: Mehmet Akgul
Project Title: Assessment of Online Review Helpfulness
MREB#: 5479

Dear Researcher(s)

The ethics application and supporting documents for MREB# 5479 entitled "Assessment of Online Review Helpfulness" have been reviewed and cleared by the MREB to ensure compliance with the Tri-Council Policy Statement and the McMaster Policies and Guidelines for Research Involving Human Participants.

The application protocol is cleared as revised without questions or requests for modification. The above named study is to be conducted in accordance with the most recent approved versions of the application and supporting documents.

If this project includes planned in-person contact with research participants, then procedures for addressing COVID-19 related risks must be addressed according to the current processes communicated by the Vice-President (Research) and your Associate Dean (Research). All necessary approvals must be secured before in-person contact with research participants can take place.

Ongoing clearance is contingent on completing the Annual Report in advance of the yearly anniversary of the original ethics clearance date: Jul/16/2022. If the Annual Report is not submitted, then ethics clearance will lapse on the expiry date and Research Finance will be notified that ethics clearance is no longer valid (TCPS, Art. 6.14).

An Amendment form must be submitted and cleared before any substantive alterations are made to the approved research protocol and documents (TCPS, Art. 6.16).

Researchers are required to report Adverse Events (i.e. an unanticipated negative consequence or result affecting participants) to the MREB secretariat and the MREB Chair as soon as possible, and no more than 3 days after the event occurs (TCPS, Art. 6.15). A privacy breach affecting participant information should also be reported to the MREB secretariat and the MREB Chair as soon as possible. The Reportable Events form is used to document adverse events, privacy breaches, protocol deviations and participant complaints.

Document Type	File Name	Date	Version
Recruiting Materials	all_recruitment_posters_updated	Jul/13/2021	2
Recruiting Materials	all_screening_questions_updated	Jul/13/2021	2
Test Instruments	all_demographic_forms_updated	Jul/13/2021	2
Test Instruments	all_instructions_forms_updated	Jul/13/2021	2
Test Instruments	all_survey_forms_updated	Jul/13/2021	2
Consent Forms	all_letter_of_information_consent_forms_updated	Jul/13/2021	2
Consent Forms	all_preamble_forms_updated	Jul/13/2021	2
Response Documents	response_to_reviewers_updated	Jul/13/2021	1

Dr. Violetta Igneski

V Igneski

Dr. Violetta Igneski, MREB Chair,
Associate Professor,
Department of Philosophy, UH-308,
905-525-9140 ext. 23462,
igneski@mcmaster.ca

Dr. Sue Becker, MREB Vice-Chair,
Professor,
Department of Psychology, Neuroscience and Behaviour, PC-312,
905-525-9140 ext. 23020,
beckers@mcmaster.ca

REFERENCES

Aakash, & Aggarwal, A. G. (2020). An Analysis of EWOM Text that Contribute to EWOM Helpfulness. In A. Kumar, M. Paprzycki, & V. K. Gunjan (Eds.), *ICDSMLA 2019. Lecture Notes in Electrical Engineering* (vol. 601, pp. 893–900). Springer.

Ab Rahman, A. F. (2005). Awareness and knowledge of epilepsy among students in a Malaysian university. *Seizure: European Journal of Epilepsy, 14*(8), 593–596.

Agnihotri, A., & Bhattacharya, S. (2016). Online review helpfulness: Role of qualitative factors. *Psychology & Marketing, 33*(11), 1006-1017.

Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk Research: Review and Recommendations. *Journal of Management, 47*(4), 823–837.

Ahmad, S. N., & Laroche, M. (2015). How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using latent semantic analysis. *International Journal of Electronic Commerce, 20*(1), 76-111.

Alzate, M., Arce-Urriza, M., & Cebollada, J. (2022). Mining the text of online consumer reviews to analyze brand image and brand positioning. *Journal of Retailing and Consumer Services, 67*, 102989.

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the pricing power of product features by mining consumer reviews. *Management science, 57*(8), 1485-1509.

Banerjee, S., & Chua, A. Y. K. (2014). A theoretical framework to identify authentic online reviews. *Online Information Review, 38*(5), 634–649.

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.

Bansal, H. S., & Voyer, P. A. (2000). Word-of-Mouth Processes within a Services Purchase Decision Context. *Journal of Service Research, 3*(2), 166–177.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology, 5*(4), 323–370.

Bayrakdar, S., Yucedag, I., Simsek, M., & Dogru, I. A. (2020). Semantic analysis on social networks: A survey. *International Journal of Communication Systems, 33*(11), e4424.

Bayus, B. L. (2013). Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community. *Management Science, 59*(1), 226–244.

Bee, C. C., & Madrigal, R. (2013). Consumer uncertainty: The influence of anticipatory emotions on ambivalence, attitudes, and intentions. *Journal of Consumer Behaviour, 12*(5), 370–381.

Berger, J. (2014). Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology, 24*(4), 586–607.

Berlo D. K. (1960). *The process of communication; an introduction to theory and practice.* Holt, Rinehart & Winston.

Bestelmeyer, P. E. G., Kotz, S. A., & Belin, P. (2017). Effects of emotional valence and arousal on the voice perception network. *Social Cognitive and Affective Neuroscience, 12*(8), 1351–1358.

Beuscart, J.-S., Mellet, K., & Trespeuch, M. (2016). Reactivity without legitimacy? Online consumer reviews in the restaurant industry. *Journal of Cultural Economy, 9*(5), 458–475.

Biswas, B., Sengupta, P., & Ganguly, B. (2021). Your reviews or mine? Exploring the determinants of "perceived helpfulness" of online reviews: a cross-cultural study. *Electron Markets, 32*(3), 1083–1102

Bjering, E., Havro, L. J., & Moen, Ø. (2015). An empirical investigation of self-selection bias and factors influencing review helpfulness. *International Journal of Business and Management, 10*(7), 16.

Boiy, E., & Moens, M.-F. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information Retrieval, 12*(5), 526–558.

Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications, 162*, 113746.

Boyd, R. L. (2017). Psychological Text Analysis in the Digital Humanities. In S. Hai-Jew (Ed.), *Data Analytics in Digital Humanities. Multimedia Systems and Applications. (*pp. 161–189). Springer.

Breckler, S. J. (1994). A comparison of numerical indexes for measuring attitude ambivalence. *Educational and psychological measurement, 54*(2), 350-365.

Brescoll, V. L., & Uhlmann, E. L. (2008). Can an Angry Woman Get Ahead?: Status Conferral, Gender, and Expression of Emotion in the Workplace. *Psychological Science, 19*(3), 268–275.

Browne, G. J., Pitts, M. G., & Wetherbe, J. C. (2007).. Cognitive Stopping Rules for Terminating Information Search in Online Tasks. *MIS Quarterly, 31*(1), 89-104.

Brynjolfsson, E., Geva, T., & Reichman, S. (2016). Crowd-Squared: Amplifying the Predictive Power of Search Trend Data. *MIS Quarterly, 40*(4), 941–961.

Bughin, J., Doogan, J., & Vetvik, O. J. (2010). A new way to measure word-of- mouth marketing. *McKinsey Quarterly, 2*(1), 113–116.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*(1), 3-25.

Cao, C., Yan, J., & Li, M. (2018). Understanding the Determinants of Online Consumer Review Helpfulness in Social Networking Service Context. *PACIS 2018 Proceedings*. 200. https://aisel.aisnet.org/pacis2018/200

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), 752–766.

Chakraborty, U., & Bhat, S. (2018). The Effects of Credible Online Reviews on Brand Equity Dimensions and Its Consequence on Consumer Behavior. *Journal of Promotion Management, 24*(1), 57–82.

Chandler, J., & Shapiro, D. (2016). Conducting Clinical Research Using Crowdsourced Convenience Samples. *Annual Review of Clinical Psychology, 12*(1), 53–81.

Chang, C. (2011). Feeling Ambivalent About Going Green. *Journal of Advertising, 40*(4), 19–32.

Chatterjee, S. (2020). Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management, 85*, 102356.

Chen, C., & Chellappa, R. (2009, May). Music sales and online user activity at social networking sites. *Fifth symposium of the Heinz College's iLab Center for Digital Media Research, Carnegie Mellon University on Statistical Challenges in Electronic Commerce Research Pittsburgh,* 30-31.

Chen, M.-Y. (2016). Can two-sided messages increase the helpfulness of online reviews? *Online Information Review, 40*(3), 316–332.

Chen, P.-Y., Dhanasobhon, S., & Smith, M. D. (2008). All Reviews are Not Created Equal: The Disaggregate Impact of Reviews and Reviewers at Amazon.Com. *SSRN Electronic Journal*. https://ssrn.com/abstract=918083

Cheng, Y.-C., & Ku, H.-Y. (2009). An investigation of the effects of reciprocal peer tutoring. *Computers in Human Behavior, 25*(1), 40–49.

Cheung, C., Sia, C.-L., & Kuan, K. (2012). Is This Review Believable? A Study of Factors Affecting the Credibility of Online Consumer Reviews from an ELM Perspective. *Journal of the Association for Information Systems, 13*(8), 618–635.

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.

Chocarro, R., Cortinas, M., & Villanueva, M. L. (2021). Different channels for different services: Information sources for services with search, experience and credence attributes. *The Service Industries Journal, 41*(3–4), 261–284.

Chou, Y.-C., Chuang, H. H.-C., & Liang, T.-P. (2022). Elaboration likelihood model, endogenous quality indicators, and online review helpfulness. *Decision Support Systems, 153*, 113683.

Chow, I. H., Lau, V. P., Wing-chun Lo, T., Sha, Z., & Yun, H. (2007). Service quality in restaurant operations in China: Decision- and experiential-oriented perspectives. *International Journal of Hospitality Management, 26*(3), 698–710.

Chua, A., & Banerjee, S. (2015). Understanding Review Helpfulness as a Function of Reviewer Reputation, Review Rating, and Review Depth. *Journal of the Association for Information Science and Technology, 66*(2), 354-362.

Chua, B.-L., Karim, S., Lee, S., & Han, H. (2020). Customer Restaurant Choice: An Empirical Analysis of Restaurant Types and Eating-Out Occasions. *International Journal of Environmental Research and Public Health, 17*(17), 6276.

Chujo, K., & Utiyama, M. (2005). Understanding the role of text length, sample size and vocabulary size in determining text coverage. *Reading in a Foreign Language, 17*(1), 1–22.

Chung, C. K., & Pennebaker, J. W. (2019). Textual analysis. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), *Measurement in Social Psychology* (1st ed., pp. 153–173). Routledge/Taylor & Francis Group.

Clare, C. J., Wright, G., Sandiford, P., & Caceres, A. P. (2018). Why should I believe this? Deciphering the qualities of a credible online customer review. *Journal of Marketing Communications, 24*(8), 823–842.

Clauss, K., & Bardeen, J. R. (2022). The interactive effect of mental contamination and cognitive fusion on anxiety. *Journal of Clinical Psychology, 78*(4), 517–525.

Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry. *Journal of Management Information Systems, 23*(2), 149–171.

Cobb, M., Brent, L. J., Buehner, T. M., Drzymala, N., & Nelson, J. J. (2014). An Alternative Front End Analysis Strategy for Complex Systems. *Army Research Inst for The Behavioral and Social Sciences*.

Connelly, B., Certo, T., Ireland, R., & Reutzel, C. (2011). Signaling Theory: A Review and Assessment. *Journal of Management, 37*(1), 39–67.

Connors, L., Mudambi, S. M., & Schuff, D. (2011). Is It the Review or the Reviewer? A Multi-Method Approach to Determine the Antecedents of Online Review Helpfulness. *2011 44th Hawaii International Conference on System Sciences*, 1–10.

Craciun, G., & Moore, K. (2019). Credibility of negative online product reviews: Reviewer gender, reputation and emotion effects. *Computers in Human Behavior, 97*, 104–115.

Craciun, G., Zhou, W., & Shan, Z. (2020). Discrete emotions effects on electronic word-of-mouth helpfulness: The moderating role of reviewer gender and contextual emotional tone. *Decision Support Systems, 130*, 113226.

Dalton, D. W. (2021). The Effects of Machiavellianism and Ethical Environment on Whistle-blowing across Low and High Moral Intensity Settings. In K. E. Karim, T. Fogarty, R. Rutledge, R. Pinsker, J. Hasseldine, C. Bailey, & T. Pitre (Eds.), *Advances in Accounting Behavioural Research* (pp. 29–49). Emerald Publishing Limited.

Deng, T. (2020). Investigating the effects of textual reviews from consumers and critics on movie sales. *Online Information Review, 44*(6), 1245–1265.

Doh, S.-J., & Hwang, J.-S. (2009). How Consumers Evaluate eWOM (Electronic Word-of-Mouth) Messages. *CyberPsychology & Behavior, 12*(2), 193–197.

Dowling, G., & Staelin, R. (1994). A Model of Perceived Risk and Intended Risk-Handling Activity. *Journal of Consumer Research, 21*(1), 119–134.

Elias, J. J., Lacetera, N., & Macis, M. (2016). Efficiency-morality trade-offs in repugnant transactions: A choice experiment. *National Bureau of Economic Research Working Paper 22632*. https://www.nber.org/system/files/working_papers/w22632/w22632.pdf.

Fan, L., & Zhang, X. (2020). The combination signaling effect of text and image on mobile phone review Helpfulness-The moderating effect of signaling environment. *Ieee Access, 8*, 122736-122746.

Fang, Z. H., & Chen, C. C. (2022). A collaborative trend prediction method using the crowdsourced wisdom of web search engines. *Data Technologies and Applications, 56*(5), 741-761.

Faraj, S., Von Krogh, G., Monteiro, E., & Lakhani, K. R. (2016). Special Section Introduction—Online Community as Space for Knowledge Flows. *Information Systems Research, 27*(4), 668–684.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.

Ferri-García, R., & Rueda, M. D. M. (2022). Variable selection in Propensity Score Adjustment to mitigate selection bias in online surveys. *Statistical Papers, 63*(6), 1829-1881.

Filieri, R. (2014). 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. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research, 58*, 46–64.

Filieri, R., McLeay, F., Tsui, B., & Lin, Z. (2018). Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services. *Information & Management, 55*(8), 956–970.

Fowler, C., Jiao, J., & Pitts, M. (2022). Frustration and ennui among Amazon MTurk workers. *Behavior Research Methods*, 1-17.

Furner, C. P., Racherla, P., & Zhen, Z. (2013). Cultural and review characteristics in the formation of trust in online product reviews: A multinational investigation. *WHICEB 2013 Proceeding*, 90. https://aisel.aisnet.org/whiceb2013/90

Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. (2004). What Do You Do When Things Go Right? The Intrapersonal and Interpersonal Benefits of Sharing Positive Events. *Journal of Personality and Social Psychology, 87*(2), 228–245.

Gerrath, M., & Usrey, B. (2016). "The review is sponsored by": The effect of sponsorship disclosure on blog loyalty. In P. Moreau, & S. Puntoni (Eds.), *Advances in consumer research* (vol. 44, p. 454-454). Association for Consumer Research.

Goldsmith, R. E., Lafferty, B. A., & Newell, S. J. (2000). The Impact of Corporate Credibility and Celebrity Credibility on Consumer Reaction to Advertisements and Brands. *Journal of Advertising, 29*(3), 43–54.

Gray, G. M., Williams, S. A., Bludevich, B., Irby, I., Chang, H., Danielson, P. D., Gonzalez, R., Snyder, C. W., Ahumada, L. M., & Chandler, N. M. (2023). Examining Implicit Bias Differences in Pediatric Surgical Fellowship Letters of Recommendation Using Natural Language Processing. *Journal of Surgical Education, 80*(4), 547–555.

Guo, B., & Zhou, S. (2017). What makes population perception of review helpfulness: an information processing perspective. *Electronic Commerce Research, 17*, 585-608.

Hales, D. (2010). *An introduction to triangulation*. UNAIDS. https://www.unaids.org/sites/default/files/sub_landing/files/10_4-Intro-to-triangulation-MEF.pdf

Han, M. C. (2021). Thumbs down on "likes"? The impact of Facebook reactions on online consumers' nonprofit engagement behavior. *International Review on Public and Nonprofit Marketing, 18*(2), 255–272.

Has, R. G., Katz, I., Rizzo, N., Bailey, J., & Moore, L. (1992). When racial ambivalence evokes negative affect, using a disguised measure of mood. *Personality and Social Psychology Bulletin, 18*(6), 786-797.

Hasan, K. M. A., Sabuj, M. S., & Afrin, Z. (2015). Opinion mining using Naïve Bayes. *2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 511–514.

Hauser, D. J., Moss, A. J., Rosenzweig, C., Jaffe, S. N., Robinson, J., & Litman, L. (2022). Evaluating CloudResearch's Approved Group as a solution for problematic data quality on MTurk. *Behavior Research Methods*.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing, 18*(1), 38–52.

Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. *Journal of Consumer Research, 17*(4), 454.

Heung, V., & Lam, T. (2003). Customer complaint behaviour towards hotel restaurant services. *International Journal of Contemporary Hospitality Management., 15*(5), 283–289.

Hill, S., Merchant, R., & Ungar, L. (2013). Lessons learned about public health from online crowd surveillance. *Big Data, 1*(3), 160-167.

Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, 102*, 1-11.

Hong, J., & Sternthal, B. (2010). The Effects of Consumer Prior Knowledge and Processing Strategies on Judgments. *Journal of Marketing Research, 47*(2), 301–311.

Hong, Y., Huang, N., Burtch, G., & Li, C. (2016). Culture, Conformity, and Emotional Suppression in Online Reviews. *Journal of the Association for Information Systems, 17*(11), 737–758.

Hsieh, J.-K., & Li, Y.-J. (2020). Will You Ever Trust the Review Website Again? The Importance of Source Credibility. *International Journal of Electronic Commerce, 24*(2), 255–275.

Hu, N., Koh, N. S., & Reddy, S. K. (2014). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision Support Systems, 57*, 42–53.

Hu, Y. H., Chen, K., & Lee, P. J. (2017). The effect of user-controllable filters on the prediction of online hotel reviews. *Information & Management, 54*(6), 728-744.

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, L.-S., Chou, Y.-J., & Lan, I.-T. (2007). Effects of Perceived Risk, Message Types, and Reading Motives on the Acceptance and Transmission of Electronic Word-of-Mouth Communication. *Contemporary Management Research, 3*(4), 299-312.

Huete-Alcocer, N. (2017). A Literature Review of Word of Mouth and Electronic Word of Mouth: Implications for Consumer Behavior. *Frontiers in Psychology, 8*(1256)

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 216–225.

Ismagilova, E., Slade, E., Rana, N. P., & Dwivedi, Y. K. (2020). The effect of characteristics of source credibility on consumer behaviour: A meta-analysis. *Journal of Retailing and Consumer Services, 53*, 101736.

Jayasingh, S., & Thiagarajan, T. (2022). Factors Influencing the Helpfulness of Online Consumer Reviews. In L. Kalinathan, P. R., M. Kanmani, & M. S. (Eds.), *Computational Intelligence in Data Science* (vol. 654, pp. 176–183). Springer.

Jensen, M. L., Averbeck, J. M., Zhang, Z., & Wright, K. B. (2013). Credibility of Anonymous Online Product Reviews: A Language Expectancy Perspective. *Journal of Management Information Systems, 30*(1), 293–324.

Jia, R., & Reich, B. H. (2013). IT service climate, antecedents and IT service quality outcomes: Some initial evidence. *The Journal of Strategic Information Systems, 22*(1), 51–69.

Jia, R., Steelman, Z., & Reich, B. H. (2017). Using Mechanical Turk Data in IS Research: Risks, Rewards, and Recommendations. *Communications of the Association for Information Systems, 41*, Article 14. https://aisel.aisnet.org/cais/vol41/iss1/14

Jo, T. (2019). Text Mining (Vol. 45). Springer International Publishing.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 363–391.

Karimi, S., & Wang, F. (2017). Online review helpfulness: Impact of reviewer profile image. *Decision Support Systems, 96*, 39-48.

Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising, 46*(1), 141-155.

Kim, B., & Velthuis, O. (2021). From reactivity to reputation management: Online consumer review systems in the restaurant industry. *Journal of Cultural Economy, 14*(6), 675–693.

Kim, J., & Gupta, P. (2012). Emotional expressions in online user reviews: How they influence consumers' product evaluations. *Journal of Business Research, 65*(7), 985–992.

Kim, J., Lee, M., Kwon, W., Park, H., & Back, K.-J. (2022). Why am I satisfied? See my reviews – Price and location matter in the restaurant industry. *International Journal of Hospitality Management, 101*, 103111.

Kim, K., Cheong, Y., & Kim, H. (2017). User-generated product reviews on the internet: The drivers and outcomes of the perceived usefulness of product reviews. *International Journal of Advertising, 36*(2), 227–245.

Kim, S. J., Maslowska, E., & Malthouse, E. C. (2018). Understanding the effects of different review features on purchase probability. *International Journal of Advertising, 37*(1), 29–53.

Klopfer, F. J., & Madden, T. M. (1980). The Middlemost Choice on Attitude Items: Ambivalence, Neutrality, or Uncertainty?. *Personality and Social Psychology Bulletin, 6*(1), 97-101.

Korachais, C., Ir, P., Macouillard, E., & Meessen, B. (2019). The impact of reimbursed user fee exemption of health centre outpatient consultations for the poor in pluralistic health systems: lessons from a quasi-experiment in two rural health districts in Cambodia. *Health Policy and Planning, 34*(10), 740-751.

Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications, 11*(3), 205-217.

Krishnamoorthy, S. (2015). Linguistic features for review helpfulness prediction. *Expert Systems with Applications, 42*(7), 3751–3759.

Krupnikov, Y., & Levine, A. S. (2014). Cross-Sample Comparisons and External Validity. *Journal of Experimental Political Science, 1*(1), 59–80.

Kwak, S. Y., Shin, M., Lee, M., & Back, K.-J. (2023). Integrating the reviewers' and readers' perceptions of negative online reviews for customer decision-making: A mixed-method approach. *International Journal of Contemporary Hospitality Management*.

Kwok, L., & Xie, K. L. (2016). Factors contributing to the helpfulness of online hotel reviews: Does manager response play a role? *International Journal of Contemporary Hospitality Management, 28*(10), 2156–2177.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time?. *Journal of Personality and Social Psychology, 81*(4), 684–696.

Lavrakas, P. (2008). *Encyclopedia of Survey Research Methods*. Sage Publications, Inc.

Le Cessie, S., Goeman, J. J., & Dekkers, O. M. (2020). Who is afraid of non-normal data? Choosing between parametric and non-parametric tests. *European Journal of Endocrinology, 182*(2), E1–E3.

Lee, K.-T., & Koo, D.-M. (2012). Effects of attribute and valence of e-WOM on message adoption: Moderating roles of subjective knowledge and regulatory focus. *Computers in Human Behavior, 28*(5), 1974–1984.

Lee, S., & Choeh, J. Y. (2018). The interactive impact of online word-of-mouth and review helpfulness on box office revenue. *Management Decision, 56*(4), 849–866.

Lee, M., & Youn, S. (2009). Electronic word of mouth (eWOM): How eWOM platforms influence consumer product judgement. *International Journal of Advertising, 28*(3), 473–499.

Li, H., Liu, H., & Zhang, Z. (2020). Online persuasion of review emotional intensity: A text mining analysis of restaurant reviews. *International Journal of Hospitality Management, 89*, 102558.

Li, H., Wang, C. R., Meng, F., & Zhang, Z. (2019). Making restaurant reviews useful and/or enjoyable? The impacts of temporal, explanatory, and sensory cues. *International Journal of Hospitality Management, 83*, 257-265.

Li, H., Zhang, Z., Meng, F., & Zhang, Z. (2019). "When you write review" matters: The interactive effect of prior online reviews and review temporal distance on consumers' restaurant evaluation. *International Journal of Contemporary Hospitality Management, 31*(3), 1273–1291.

Li, J., Xu, X., & Ngai, E. W. (2021). Does certainty tone matter? Effects of review certainty, reviewer characteristics, and organizational niche width on review usefulness. *Information & Management, 58*(8), 103549.

Li, M.-Q., Lee, Y.-C., & Wu, L.-L. (2020). Sidedness Effect of Online Review. *PACIS 2020 Proceedings*. 125. https://aisel.aisnet.org/pacis2020/125.

Liang, S., Schuckert, M., & Law, R. (2019). How to improve the stated helpfulness of hotel reviews? A multilevel approach. *International Journal of Contemporary Hospitality Management, 31*(2), 953-977.

Licata, J., & Frankwick, G. L. (1996). University marketing: a professional service organization perspective. *Journal of Marketing for Higher Education, 7*(2), 1-16.

Lima, M., & Fernandes, T. (2015). Relationship Bonds and Customer Loyalty: A Study Across Different Service Contexts. *Proceedings of the International Conference on Exploring Services Science, Porto, Portugal, 4–6 February 2015*, 326–339.

Lin, M. S., Sharma, A., & Ouyang, Y. (2020). Role of Signals In Consumers' Economic Valuation of Restaurant Choices. *Journal of Hospitality & Tourism Research, 44*(7), 1080–1100.

Litman, L., Moss, A., Rosenzweig, C., & Robinson, J. (2021). Reply to MTurk, Prolific or panels? Choosing the right audience for online research. *SSRN Electronic Journal*. https://ssrn.com/abstract=3775075

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management, 29*(3), 458–468.

Liu, F., Li, J., Mizerski, D., & Soh, H. (2012). Self-congruity, brand attitude, and brand loyalty: A study on luxury brands. *European Journal of Marketing, 46*(7/8), 922–937.

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism management, 47*, 140-151.

Lopes, A. I., Dens, N., De Pelsmacker, P., & De Keyzer, F. (2020). Which cues influence the perceived usefulness and credibility of an online review? A conjoint analysis. *Online Information Review, 45*(1), 1–20.

Lopez, A., & Garza, R. (2022). Do sensory reviews make more sense? The mediation of objective perception in online review helpfulness. *Journal of Research in Interactive Marketing, 16*(3), 438–456.

Luo, Y., Tang, L. (Rebecca), Kim, E., & Wang, X. (2020). Finding the reviews on yelp that actually matter to me: Innovative approach of improving recommender systems. *International Journal of Hospitality Management, 91*, 102697.

Lutz, R. J., & Reilly, P. J. (1974). An Exploration of the Effects of Perceived Social and Performance Risk on Consumer Information Acquisition. *Advances in Consumer Research, 1*, 393-405. https://www.acrwebsite.org/volumes/5672/volumes/v01/NA-01/full.

Malhotra, N. K. (1984). Reflections on the Information Overload Paradigm in Consumer Decision Making. *Journal of Consumer Research, 10*(4), 436–440.

Mariani, M., & Borghi, M. (2021). Are environmental-related online reviews more helpful? A big data analytics approach. *International Journal of Contemporary Hospitality Management, 33*(6), 2065-2090.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behavior Research Methods, 44*(1), 1–23.

Martin, K., & Nissenbaum, H. (2017). Privacy interests in public records: An empirical investigation. *Harvard Journal of Law and Technology, 31*(1), 111–143.

McGraw, K. M., Hasecke, E., & Conger, K. (2003). Ambivalence, uncertainty, and processes of candidate evaluation. *Political Psychology, 24*(3), 421-448.

Meta. (2023, February 1). *Meta Reports Fourth Quarter and Full Year 2022 Results*. https://investor.fb.com/investor-news/press-release-details/2023/Meta-Reports-Fourth-Quarter-and-Full-Year-2022-Results/default.aspx

Mitra, K., Reiss, M. C., & Capella, L. M. (1999). An examination of perceived risk, information search and behavioral intentions in search, experience and credence services. *Journal of Services Marketing, 13*(3), 208–228.

Mudambi & Schuff. (2010). Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly, 34*(1), 185-200.

Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The Generalizability of Survey Experiments. *Journal of Experimental Political Science, 2*(2), 109–138.

Muralidharan, S., Yoon, H. J., Sung, Y., Miller, J., & Lee, A. (2017). Following the breadcrumbs: An analysis of online product review characteristics by online shoppers. *Journal of Marketing Communications, 23*(2), 113–134.

Myors, B., Murphy, K. R., & Wolach, A. (2008). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests.* Routledge.

Nam, K., Baker, J., Ahmad, N., & Goo, J. (2020). Determinants of writing positive and negative electronic word-of-mouth: Empirical evidence for two types of expectation confirmation. *Decision Support Systems, 129*, 113168.

Necka, E. A., Cacioppo, S., Norman, G. J., & Cacioppo, J. T. (2016). Measuring the Prevalence of Problematic Respondent Behaviors among MTurk, Campus, and Community Participants. *PLoS ONE, 11*(6), e0157732.

Newman, I. (1980). Repeated Measures: Suggested Approach for Simultaneous Independent and Dependent Measures. *Annual meeting of the American Psychological Association*. https://eric.ed.gov/?id=ED206643.

Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising, 19*(3), 39–52.

Ortony, A., Clore, G., & Collins, A. (2022). *The Cognitive Structure of Emotions*. Cambridge University Press.

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.

Pantelidis, I. S. (2010). Electronic Meal Experience: A Content Analysis of Online Restaurant Comments. *Cornell Hospitality Quarterly, 51*(4), 483–491.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*(5), 411–419.

Park, C. W., & Seo, D. R. (2018). Sentiment analysis of Twitter corpus related to artificial intelligence assistants. *2018 5th International Conference on Industrial Engineering and Applications (ICIEA)*, 495–498.

Park, J., Yi, Y., & Kang, D. (2019). The Effects of One-Sided vs. Two-Sided Review Valence on Electronic Word of Mouth (e-WOM): The Moderating Role of Sponsorship Presence. *Asia Marketing Journal, 21*(2), 1–19.

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research, 50*, 67–83.

Pavlou, P., Liang, H., & Xue, Y. (2007). Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. *MIS Quarterly, 31*(1), 105–136.

Pentina, I., Bailey, A. A., & Zhang, L. (2018). Exploring effects of source similarity, message valence, and receiver regulatory focus on yelp review persuasiveness and purchase intentions. *Journal of Marketing Communications, 24*(2), 125–145.

Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research, 67*(5), 1035-1041.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and Persuasion: Classic and Contemporary Approach*. W.C. Brown Co. Publishers.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource-dependency perspective*. Harper & Row.

Picher Vera, D., Martínez María-Dolores, S. M., & Bernal García, J. J. (2016, June). Social networks as a communication, sales and customer service tool. Analysis and effectiveness of eWOM as a marketing strategy. *14th International Conference on Marketing*.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of personality and social psychology, 71*(3), 431.

Qahri-Saremi, H., & Montazemi, A. (2019). Factors Affecting the Adoption of an Electronic Word of Mouth Message: A Meta-Analysis. *Journal of Management Information Systems, 36*(3), 969–1001.

Qahri-Saremi, H., & Montazemi, A. R. (2023). Negativity bias in the diagnosticity of online review content: the effects of consumers' prior experience and need for cognition. *European Journal of Information Systems, 32*(4), 717-734.

Quaschning, S., Pandelaere, M., & Vermeir, I. (2015). When Consistency Matters: The Effect of Valence Consistency on Review Helpfulness. *Journal of Computer-Mediated Communication, 20*(2), 136–152.

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.

Racherla, P., Furner, C., & Babb, J. (2012). Conceptualizing the Implications of Mobile App Usage and Stickiness: A Research Agenda. *SSRN Electronic Journal.* https://ssrn.com/abstract=2187056

Raghunathan, R., & Pham, M. T. (1999). All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making. *Organizational Behavior and Human Decision Processes, 79*(1), 56–77.

Razzouk, R., & Johnson, T. (2012). Shared Cognition. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 3056–3058). Springer.

Reich, T., Kaju, A., & Maglio, S. J. (2023). How to overcome algorithm aversion: Learning from mistakes. *Journal of Consumer Psychology, 33*(2), 285–302.

Ren, G., & Hong, T. (2019). Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews. *Information Processing & Management, 56*(4), 1425-1438.

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM, 43*(12), 45-48.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review, 1*(1), 60–85.

Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research, 12*(2), 479–522.

Rothman, N., Pratt, M., Rees, L., & Vogus, T. (2017). Understanding the Dual Nature of Ambivalence: Why and When Ambivalence Leads to Good and Bad Outcomes. *The Academy of Management Annals, 11*(1), 33–72.

Russell, C. A., Russell, D. W., & Klein, J. (2011). Ambivalence toward a country and consumers' willingness to buy emblematic brands: The differential predictive validity of objective and subjective ambivalence measures on behavior. *Marketing Letters, 22*(4), 357–371.

Sainani, K. L. (2012). Dealing With Non-normal Data. *PM&R, 4*(12), 1001–1005.

Schiffman, L. G., & Kanuk, L. L. (2000). *Consumer behavior (7th ed.)*. Prentice Hall.

Septianto, F., Kemper, J. A., & Choi, J. (Jane). (2020). The power of beauty? The interactive effects of awe and online reviews on purchase intentions. *Journal of Retailing and Consumer Services, 54*, 102066.

Shafir, T. (2015). Movement-based strategies for emotion regulation. In M. L. Bryant (Ed.), *Handbook on emotion regulation: Processes, cognitive effects and social consequences*, 231–249.

Shaft, T., Tian, A. C., & Whang, S. Y. (2020). Anxious and Angry: A Replication Investigating the Effects of Emotions on Perceptions of Online Review Helpfulness. *AIS Transactions on Replication Research, 6*(1), 17. https://aisel.aisnet.org/trr/vol6/iss1/17

Sharma, P., Chen, I. S. N., & Luk, S. T. K. (2012). Gender and age as moderators in the service evaluation process. *Journal of Services Marketing, 26*(2), 102–114.

Shin, S. Y., Van Der Heide, B., Beyea, D., Dai, Y. (Nancy), & Prchal, B. (2017). Investigating moderating roles of goals, reviewer similarity, and self-disclosure on the effect of argument quality of online consumer reviews on attitude formation. *Computers in Human Behavior, 76*, 218–226.

Siddiqi, U. I., Sun, J., & Akhtar, N. (2020). The role of conflicting online reviews in consumers' attitude ambivalence. *The Service Industries Journal, 40*(13–14), 1003–1030.

Simon, H. (1957) *Models of Man, Social and* Rational. Wiley.

Smirnov, O., & Hsieh, P.-H. (2022). COVID-19, climate change, and the finite pool of worry in 2019 to 2021 Twitter discussions. *Proceedings of the National Academy of Sciences, 119*(43), e2210988119.

Snyder, A. I., & Tormala, Z. L. (2017). Valence asymmetries in attitude ambivalence. *Journal of Personality and Social Psychology, 112*(4), 555–576.

Soror, A. A., Hammer, B. I., Steelman, Z. R., Davis, F. D., & Limayem, M. M. (2015). Good habits gone bad: Explaining negative consequences associated with the use of mobile phones from a dual-systems perspective: Good habits gone bad. *Information Systems Journal, 25*(4), 403–427.

Souki, G. Q., Oliveira, A. S. de, Guerreiro, M. M. M., Mendes, J. da C., & Moura, L. R. C. (2022). Do memorable restaurant experiences affect eWOM? The moderating effect of consumers' behavioural engagement on social networking sites. *The TQM Journal*.

Sousa, R. F. D., & Pardo, T. A. S. (2021). The Challenges of Modeling and Predicting Online Review Helpfulness. *Anais Do XVIII Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2021)*, 727–738.

Srivastava, V., & Kalro, A. D. (2019). Enhancing the Helpfulness of Online Consumer Reviews: The Role of Latent (Content) Factors. *Journal of Interactive Marketing, 48*(1), 33–50.

Steelman, Z. R., Hammer, B. I., & Limayem, M. (2014). Data collection in the digital age: innovative alternatives to student samples. *MIS quarterly, 38*(2), 355–378.

Sun, T., Youn, S., Wu, G., & Kuntaraporn, M. (2006). Online Word-of-Mouth (or Mouse): An Exploration of Its Antecedents and Consequences. *Journal of Computer-Mediated Communication, 11*(4), 1104–1127.

Sun, X., Han, M., & Feng, J. (2019). Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems, 124*, 113099.

Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences, 109*(21), 8038–8043.

Tang, C., & Guo, L. (2015). Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication. *Marketing Letters, 26*(1), 67–80.

Thapa, S., Guzmán, F., & Paswan, A. K. (2022). How isolation leads to purchasing luxury brands: The moderating effects of COVID-19 anxiety and social capital. *Journal of Product & Brand Management, 31*(6), 984–1001.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 361-386). Lawrence Erlbaum Associates, Inc.

Ullah, R., Zeb, A., & Kim, W. (2015). The impact of emotions on the helpfulness of movie reviews. *Journal of applied research and technology, 13*(3), 359-363.

Ullrich, J. (2012). A multivariate approach to ambivalence: It is more than meets the IV. In J. I. Krueger (Ed.), *Social judgment and decision making* (pp. 115–132). Psychology Press.

Uribe, R., Buzeta, C., & Velásquez, M. (2016). Sidedness, commercial intent and expertise in blog advertising. *Journal of Business Research, 69*(10), 4403–4410.

Van Harreveld, F., Nohlen, H. U., & Schneider, I. K. (2015). The ABC of ambivalence: Affective, behavioral, and cognitive consequences of attitudinal conflict. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (vol. 52, pp. 285–324). Academic Press.

Van Harreveld, F., Pligt, J., Vries, N. K., Wenneker, C., & Verhue, D. (2004). Ambivalence and information integration in attitudinal judgment. *British Journal of Social Psychology, 43*(3), 431–447.

Van Harreveld, F., Rutjens, B. T., Rotteveel, M., Nordgren, L. F., & Van Der Pligt, J. (2009). Ambivalence and decisional conflict as a cause of psychological discomfort: Feeling tense before jumping off the fence. *Journal of Experimental Social Psychology, 45*(1), 167–173.

Wang, C. C., Li, M. Z., & Yang, Y. Y. (2015). Perceived usefulness of word-of-mouth: An analysis of sentimentality in product reviews. In Wang, L., Uesugi, S., Ting, IH., Okuhara, K., Wang, K. (Eds), *Multidisciplinary Social Networks Research. MISNC 2015. Communications in Computer and Information Science* (vol 540, pp. 448-459). Springer.

Wang, F., & Karimi, S. (2019). This product works well (for me): The impact of first-person singular pronouns on online review helpfulness. *Journal of Business Research, 104*, 283-294.

Wang, J.-N., Du, J., & Chiu, Y.-L. (2020). Can online user reviews be more helpful? Evaluating and improving ranking approaches. *Information & Management, 57*(8), 103281.

Wang, S., Cunningham, N., & Eastin, M. (2015). The Impact of eWOM Message Characteristics on the Perceived Effectiveness of Online Consumer Reviews. *Journal of Interactive Advertising, 15*(2), 151–159.

Wang, X., Tang, L. R., & Kim, E. (2019). More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness?. *International Journal of Hospitality Management, 77*, 438-447.

Wang, Y., Kim, J., & Kim, J. (2021). The financial impact of online customer reviews in the restaurant industry: A moderating effect of brand equity. *International Journal of Hospitality Management, 95*, 102895.

Wang, Y., Wang, J., & Yao, T. (2019). What makes a helpful online review? A meta-analysis of review characteristics. *Electronic Commerce Research, 19*, 257-284.

Weick, K.E. (1995). *Sensemaking in organizations*. Sage Publications, Inc.

Weiss, S. M., Indurkhya, N., & Zhang, T. (2015). *Fundamentals of Predictive Text Mining*. Springer.

Weng, J., & DeMarree, K. G. (2019). An Examination of Whether Mindfulness Can Predict the Relationship Between Objective and Subjective Attitudinal Ambivalence. *Frontiers in Psychology, 10*, 854.

Westbrook, R. A. (1987). Product/Consumption-Based Affective Responses and Postpurchase Processes. *Journal of marketing research, 24*(3), 258-270.

Wetzer, I. M., Zeelenberg, M., & Pieters, R. (2007). “Never eat in that restaurant, I did!”: Exploring why people engage in negative word-of-mouth communication. *Psychology and Marketing, 24*(8), 661–680.

Whitler, K. A. (2014, July 17). *Why Word of Mouth Marketing Is The Most Important Social Media*. Forbes.https://www.forbes.com/sites/kimberlywhitler/2014/07/17/why-word-of-mouth-marketing-is-the-most-important-social-media/

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.

Woo, S. E., Keith, M., & Thornton, M. A. (2015). Amazon Mechanical Turk for Industrial and Organizational Psychology: Advantages, Challenges, and Practical Recommendations. *Industrial and Organizational Psychology, 8*(2), 171–179.

Xiao, B., & Benbasat, I. (2007). E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact. *MIS Quarterly, 31*(1), 137-209.

Xu, C., Zheng, X., & Yang, F. (2023). Examining the effects of negative emotions on review helpfulness: The moderating role of product price. *Computers in Human Behavior, 139*, 107501.

Xu, D., Ye, Q., Hong, H., & Sun, F. (2022). Emotions for attention in online consumer reviews: the moderated mediating role of review helpfulness. I*ndustrial Management & Data Systems, 122*(3), 729-751.

Yan, T., Kumar, V., & Ganesan, D. (2010). CrowdSearch: Exploiting crowds for accurate real-time image search on mobile phones. *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, 77–90.

Yang, L., & Unnava, H. (2016). Ambivalence, Selective Exposure, and Negativity Effect. *Psychology & Marketing, 33*(5), 331–343.

Yang, S., Yao, J., & Qazi, A. (2020). Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining. *Information Processing & Management, 57*(2), 102179.

Yelp. (2023). *Yelp Fast Facts*. https://www.yelp-press.com/company/fast-facts/default.aspx

Yin, D., Bond, S. D., & 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., Bond, S. D., & Zhang, H. (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.

Yin, D., de Vreede, T., Steele, L., & de Vreede, G.-J. (2021). Cross-Review Incoherence and Purchase Deferral. *International Conference on Information Systems 2021 Proceedings.* 2.https://aisel.aisnet.org/icis2021/digital_commerce/digital_commerce/2.

Yin, G., Zhang, Q., & Li, Y. (2014). Effects of emotional valence and arousal on consumer perceptions of online review helpfulness. *Americas Conference on Information Systems 2014 Proceedings*. 6. https://aisel.aisnet.org/amcis2014/Posters/SocialComputing/6

Yoo, K.-H., & Gretzel, U. (2009). Comparison of Deceptive and Truthful Travel Reviews. In W. Höpken, U. Gretzel, & R. Law (Eds.), In*formation and Communication Technologies in Tourism 2009* (pp. 37–47). Springer.

Zech, E., & Rimé, B. (2005). Is talking about an emotional experience helpful? Effects on emotional recovery and perceived benefits. *Clinical Psychology & Psychotherapy, 12*(4), 270–287.

Zhang, H., Sun, J., Liu, F., & G. Knight, J. (2014). Be rational or be emotional: Advertising appeals, service types and consumer responses. *European Journal of Marketing, 48*(11/12), 2105-2126.

Zhang, H., Wang, S., Chen, T.-H. (Peter), & Hassan, A. E. (2021). Are Comments on Stack Overflow Well Organized for Easy Retrieval by Developers?. *ACM Transactions on Software Engineering and Methodology, 30*(2), 1–31.

Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research, 63*(12), 1336-1341.

Zhang, K. Z. K., Lee, M. K. O., & Zhao, S. J. (2010). Understanding the informational social influence of online review platforms. *International Conference on Information Systems 2010 Proceedings*. 71. https://aisel.aisnet.org/icis2010_submissions/71

Zhang, P., Lee, H.-M., Zhao, K., & Shah, V. (2019). An empirical investigation of eWOM and used video game trading: The moderation effects of product features. *Decision Support Systems, 123*, 113076.

Zhang, T., Lee, B., Zhu, Q., Han, X., & Ye, E. M. (2020). Multi-dimension topic mining based on hierarchical semantic graph model. *IEEE Access, 8*, 64820-64835.

Zhang, Z., Wang, X., & Wu, R. (2021). Is the devil in the details? Construal-level effects on perceived usefulness of online reviews for experience services. *Electronic Commerce Research and Applications, 46*, 101033.

Zhang, Z., Zhang, Z., & Law, R. (2014). Positive and Negative Word of Mouth about Restaurants: Exploring the Asymmetric Impact of the Performance of Attributes. Asia Pacific *Journal of Tourism Research, 19*(2), 162–180.

Zhao, C. (2020). Are Negative Sentiments "Negative" for Review Helpfulness?. *Pacific Conference on Information Systems 2020 Proceedings*. 22. https://aisel.aisnet.org/pacis2020/22

Zhou, S., & Guo, B. (2015). The Interactive Effect of Review Rating and Text Sentiment on Review Helpfulness. In H. Stuckenschmidt & D. Jannach (Eds.), *E-Commerce and Web Technologies. EC-Web 2015. Lecture Notes in Business Information Processing* (vol. 239, pp. 100–111). Springer.

Zhou, S., & Guo, B. (2017). The order effect on online review helpfulness: A social influence perspective. *Decision Support Systems, 93*, 77-87.

Zhou, Y., & Yang, S. (2019). Roles of review numerical and textual characteristics on review helpfulness across three different types of reviews. *IEEE Access, 7*, 27769-27780.

Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research, 15*(4), 267.

Zhu, X. (Susan), Barnes-Farrell, J. L., & Dalal, D. K. (2015). Stop Apologizing for Your Samples, Start Embracing Them. *Industrial and Organizational Psychology, 8*(2), 228–232.