The Impact of Online Reviews on Consumer Evaluations and Decision Making: An Analysis of Review Volume and User-Generated Photos

Matheus Lima de Castro Menezes

Imperial College Business School

3rd of May 2023

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at Imperial College London, May 2023

Declaration of Originality

I hereby declare that this thesis is my own original work and has not been submitted, in whole or in part, to any other institution for any degree or qualification. All sources consulted have been appropriately acknowledged, and all direct quotations and paraphrases are properly cited.

Copywrite Declaration

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr Rajesh Bhargave, for his unwavering support, guidance, and encouragement throughout my doctoral journey. His expertise, patience, and mentorship has been invaluable in helping me develop my research and writing skills, and I am forever grateful for the opportunity to learn from him. Rajesh has not only been an outstanding advisor but also, I hope to have found a lifelong friend in him.

I would like to thank Dr Bryan Usrey for his dedication and guidance as colleague and friend. I would also like to extend my gratefulness to the PhD thesis examiners, Professor Eduardo Andrade and Dr Johannes Hattula, for their useful comments and feedback.

I would like to extend my appreciation to my friends: Nick Shields, Sophie Barthel, Antonio Luiz Gomes, and Stefano Benigni. Their encouragement, friendship, and constant reminders of the importance of maintaining a healthy balance between work and life has been instrumental in overcoming challenges and fostering personal growth. Our long conversations and shared experiences have been an invaluable source of motivation and inspiration.

Additionally, I would like to express my gratitude to my fellow doctoral candidates: Barbara Duffek, Tanita Yönel, and Yanting He. They have provided invaluable intellectual stimulation and motivation during my time at Imperial Business School. Our discussions have fostered a supportive environment in which we have all been able to thrive academically and personally. I am truly fortunate to have had the opportunity to work alongside such talented individuals and look forward to maintaining these connections in the years to come.

My most profound gratitude goes to my loving family, who have been my pillar of support throughout this challenging and rewarding journey, further intensified by the added difficulties posed by the COVID-19 pandemic.

To my sister, thank you for your constant love, understanding, and companionship during the ups and downs of this journey. Your unwavering faith in me and your willingness to lend a listening ear or a helping hand whenever I needed it has been a source of immense comfort and strength, especially during these trying times. To my brother-in-law, who has not only served as an example of academic excellence and mental strength but also a friend to count on. His encouragement and belief in my abilities have significantly contributed to my success, and I am incredibly grateful for his presence in my life.

Finally, I would like to extend a special thank you to my mother and father, whose fearlessness, guidance, and nurturing have instilled in me the values of hard work, resilience, and the pursuit of excellence. They have taught me the importance of fortitude, even in the face of unprecedented challenges personal and professional. Their unwavering belief in my potential and their steadfast encouragement have been the driving force behind my perseverance and determination to succeed in my doctoral studies. Eu amo vocês.

To all those who have contributed to my academic journey, directly or indirectly, I am eternally grateful.

Abstract

This thesis investigates the impact of online social influence on consumer behaviour, specifically within the context of online reviews. It examines how review volume and user-generated photos affect consumer evaluations and decision-making. In Chapter 2, I introduce a novel phenomenon, the N-effect, which explores how opinion volume influences the content of online evaluations. I find that as the number of opinions increases, the content becomes more emotional and less analytical. In Chapter 3, I investigate the role of user-generated photos in shaping purchase intentions. This research demonstrates that photos can enhance review helpfulness, even when they lack diagnostic information. This effect is driven by the confidence signalled by the reviewer when posting a review with a photo, which is later assimilated by readers, leading to increased perceived helpfulness and purchase likelihood.

This thesis makes several theoretical and practical contributions to the literature on human interaction with technology. Theoretically, it expands our understanding of online social influence by examining the dynamics of online opinion expression and content. I contribute to the literature on group size by demonstrating how responsibility may be lost in online contexts. Furthermore, the findings provide insights into the social influence of photos on viewers and the role of pseudo-evidence in shaping beliefs and attitudes.

From a practical standpoint, this research offers valuable insights for online platform managers and marketers on interpreting and using consumer-written reviews. Overall, this thesis contributes to the existing literature on online social influence and provides insights for businesses to improve communication and interpretation with consumers by better understanding and leveraging online reviews and opinions.

Keywords: social influence, judgement and decision making, online opinions, user-generated photos

Table of Contents

Chapter 1: Introduction

Caveat Emptor

Before beginning my doctoral journey, I encountered a publication in the New England Journal of Medicine titled "The Collective Dynamics of Smoking in a Large Social Network" by Nicholas A. Christakis and James H. Fowler (2008). In this work, the authors examined the spread of smoking behaviour among individuals within a social network. They discovered that the decision to quit smoking is not made independently by an individual, but rather is influenced by collective pressures within the network, with effects extending up to three degrees of separation. They also observed that networks become progressively more polarised, with relatively fewer social ties between smokers and non-smokers. Beyond the implications for clinical and public health interventions, this study highlights an intriguing phenomenon: individuals' ability to influence one another without physical presence.

This idea that individuals can influence each other's behaviour only by belonging to the same network resonated with me. In the past decade, consumer judgment and decision-making have undergone an unprecedented digital transformation. Websites displaying reviews observed an increase in conversion rates across multiple product categories. For example, the effect of review sites on increasing conversion rates was 38% in home appliances and electronics, by 13.8% in food and beverages, and 15% in books, home and gardening, and, health and beauty products in 2022[1]. If social influence can spread through a group in an offline setting without physical presence, how are behaviours influenced in an online environment where the absence of physical presence is a defining factor? Moreover, how do such influencing factors affect the aggregate opinion?

Social influence plays a critical role in shaping consumer behaviour (Festinger, 1954), particularly in online contexts where opinions are readily available and often serve as informational signals for consumers (Katona et al., 2011; Zhang et al., 2014). In this regard, the concept of the wisdom of crowds—the notion that a group's collective intelligence surpasses that of any individual member—gains prominence in online settings (Frey & van de Rijt, 2021). Sharing opinions and evaluations of products and services can offer valuable information to consumers.

The strategy employed by consumers to bridge the knowledge gap left by marketers has been the focus of numerous research studies. The consensus highlights the influential effect that

[1] https://www-statista-com.iclibezp1.cc.ic.ac.uk/statistics/1322695/online-reviews-conversion-rates-growth-by-category/

aggregated accounts can have on consumer purchasing decisions online. Consumers interpret aggregated opinions as a risk mitigator when making choices. A collection of individual opinions is considered wise, however, gaps remain in our understanding of how the collection process of these opinions is conducted and whether they are free from bias.

For instance, in my first work, I demonstrate that in this sequential process, later opinions differ from earlier ones, a condition that undermines the reliability of the wisdom of crowds. That is to say, an additional opinion does not simply increase the amount of information in a consistent manner but alters the characteristics of the information provided. Furthermore, as part of this thesis, I reveal that much of the online review content, such as consumer-uploaded photos, does not solely serve as an indicator of product quality, but instead functions as a signal of reviewers' confidence. These findings, which I discuss in greater detail in Chapter 2 and Chapter 3, suggest that consumers should be cautious of how this accumulation can influence their decisions and skew the decision-making process of the group.

Given these complexities, *caveat emptor* – "let the buyer beware", underscores the need for consumers to critically evaluate online information before making decisions. Understanding the dynamics of *online social influence* and its effects on consumer information processing is essential, particularly in the context of online reviews. This thesis aims to contribute to this understanding by investigating the impact of online social influence and information processing on consumer behaviour. Through a series of studies, this research examines the influence of review volume and user-generated photos on consumer evaluations and decision-making, offering insights into how online evaluations can affect consumer behaviour and be leveraged to drive desired outcomes.

Caveat Lector

Caveat lector, or "let the reader beware", sets the stage for understanding the current landscape of online reviews. Today, some of the most popular products, services, and experiences have accumulated tens of thousands of opinions on review platforms. For example, Yelp has had a 9% year on year increase in the number of reviews, totalling 244 million at the end of 2021[2] and Trip Advisor has accumulated more than 1 billion user reviews[3]. Yet little is known about how the sheer volume of reviews already posted affects the contributions of those posting a review later on. Moreover, managers are unclear how individual opinions from vast accumulations should be interpreted. In this first work (Chapter 2), I investigate the dynamics of how people express their opinion in response to the growing volume of opinions by others for a given target. I discover a novel phenomenon, dubbed the N-effect, which influences how online evaluations are contributed.

Analyses of two datasets covering 75 restaurants & hotels across 20,634 opinion texts and four experimental studies reveal that as the volume of opinions (N) increase, opinion content becomes more emotional and less analytical for a given target. Results further rule out any conformity process and show that online contributors interpret N as a cue for how much responsibility towards readers they should feel when sharing their opinion online. I discuss the implications for existing research on online reviews and how practitioners should react to such volumes.

This work makes several theoretical contributions. It adds to the literature on online social influence and opinion sharing by examining how the volume of online opinions influences contributor content. It is the first study to examine the impact of review accumulation on online opinions and highlight the influence of post-purchase written evaluations on readers. The work also provides evidence of how a reduced sense of responsibility may lead to affect-rich content and extends group size literature by demonstrating how responsibility may be lost in online contexts. The findings highlight the need for a deeper understanding of the dynamics of online opinion expression and content.

This work also make several practical contributions. As online review volumes increase, it's important for businesses to have strategies to address the potential impact this accumulation may have on opinion collection. Marketers and online platforms should balance analytical and emotional opinions in online reviews, as past research has found that detailed information is more persuasive in certain contexts, while emotions can enhance persuasiveness more broadly. Furthermore, the value of a review depends on the product being reviewed, with analytical evaluations suitable for

[2] https://www-statista-com.iclibezp1.cc.ic.ac.uk/statistics/278032/cumulative-number-of-reviews-submitted-to-yelp/
[3] https://www-statista-com.iclibezp1.cc.ic.ac.uk/statistics/684862/tripadvisor-number-of-reviews/

making improvements and comparisons, while emotional evaluations can be more helpful for evaluating ambiguous criteria (Baek et al., 2012; Huang et al., 2013). Alternatively, review value may also depend on consumer level of expertise (Park & Kim, 2008). This research provides practical insights for online platform managers on how to interpret and utilize consumer-written reviews, highlighting the potential biases that can arise from misinterpreting or ignoring warnings and distorted perceptions.

Caveat Spectator

Consumers perceive reviews of product and services that include photos as particularly helpful (Ceylan et al., 2023). Review helpfulness is important because consumers are more likely to use such reviews in their decisions (Hong et al., 2017). In this second work (Chapter 3) I test the effect of including a photo with a review and its impact on review helpfulness using reviews and photos from a wide variety of domains including restaurants, products, and tourist experiences. I test this effect across a number of studies including field data from an online review platform and four experiments.

Building upon the notion of *caveat lector*, introduced in the previous section, we now turn our attention to *caveat spectator*, or "let the viewer beware". Consumers perceive reviews of products and services that include photos as particularly helpful or useful (An et al., 2020). On many platforms, review helpfulness is often termed 'useful'. Votes of usefulness are crucial for review platforms, as they help identify and promote high-quality content, leading to a better user experience and increased trust in the platform (Zinko et al., 2021).

In this second work (Chapter 3), I test the effect of including a photo with a review and its impact on review helpfulness or usefulness, using reviews and photos from a wide variety of domains, including restaurants, products, and tourist experiences. I examine this effect across a number of studies, encompassing field data from an online review platform and four experiments. This exploration serves to remind readers and viewers alike to exercise caution when interpreting both textual and visual information presented in online reviews, and underscores the importance of usefulness votes in shaping the overall credibility and success of review platforms.

These studies show that photos can increase review helpfulness even when the photo provides no diagnostic information, such as photos that do not indicate valence and those that have minimal or no correspondence with the review text. Further, I argue that the effect of photos on the helpfulness of the review is driven by the confidence signalled by the reviewer in posting a review and that confidence later being assimilated by the readers of the review to increase their perceived helpfulness and purchase likelihood.

This work makes several theoretical contributions. It expands on the existing online review literature by emphasising the perceived value of user-generated photos on review platforms. Even when these images do not possess clear diagnostic value. I also contribute to signalling literature by showing that photos act as visual cues for reviewer confidence and generate pseudo-evidence, leading to increased reader confidence and review value. These findings offer insights into the social influence of photos on viewers, going beyond review value or purchase intent, and suggests that consumers may not rely solely on review information. The findings extend previous research on information acquisition by showing that review value can be influenced by perceived reviewer confidence. I suggest that the photo heuristic generates pseudo-evidence, an illusion that affects beliefs and attitudes despite lacking evidence in support of the claims made.

This work also makes several practical contributions. The findings indicate that user-generated photos serve as more than visual information. This enhances the value of review platforms by boosting the perceived usefulness of individual reviews and elevating the overall perception of the platform. These results align with past literature on photos' impact on trust (Newman et al., 2012), recall (Cardwell et al., 2016), and product evaluations (Mantonakis et al. 2014). However, photos may also pose risks to consumers by leading to false claims (Cardwell et al. 2016) or misinformation (Fenn et al., 2019), potentially resulting in false confidence or unsupported beliefs.

In summary, this thesis aims to inform platform designers and marketers on how to leverage the wisdom of crowds in online settings while also acknowledging the potential for alternative explanations. By providing insights into the dynamics of online social influence and information processing, this research contributes to the growing body of literature on human interaction with technology and informs future research in this area.

The remainder of this thesis is organized as follows: Chapter 2 delves into the N-effect, which explores how the volume of opinions influences the content of online evaluations. This chapter includes two field data studies and four experimental studies that examine the effect of opinion volume on online opinions. In Chapter 3, I investigate the role of user-generated photos in shaping purchase intentions. This chapter examines the value of photos on review platforms and how they influence review value and consumer confidence. Finally, in the general discussion, I provide an overview of the key findings and their implications for online platform managers and marketers. Overall, this thesis aims to contribute to the existing literature on online social influence and provide practical insights into how businesses can better understand and leverage online reviews and opinions to improve their products and services.

Chapter 2: The N-Effect: Examining the Dynamics of Online Opinion Expression in Response to Volume

Caveat Lector

Introduction

Online reviews have become a significant source of information for consumers. With the proliferation of online platforms, the number of reviews is often perceived as a reliable indicator of the quality of information about a product or service. For example, popular tourist attractions like the Eiffel Tower and Buckingham Palace have, respectively, 320,758 and 148,183 reviews on Google. Online platforms often equate the volume of reviews with the overall quality of the information, Moreover, marketers often assume that the influence or biases in the collection of reviews is mitigated by their quantity. However, this may not always be true for online opinions.

I present a new phenomenon that demonstrates how online opinions are influenced by the accumulation of other opinions. I term this the "N-effect," where "N" refers to a contributor's ordinal position in a particular series of opinions. I find that contributors judge targets differently depending on their N, suggesting that N is an imperfect signal of information reliability in online settings. Later opinions are constructed differently from initial ones. More specifically, consumers provide less analytical and more emotional opinions as N increases.

As an example, consider the case of Adriana posting a review of the Eiffel Tower. Online platforms have been collecting reviews of this historical landmark from consumers for some time. Suppose that Adriana was one of the first people to post a review on one of these platforms. When evaluating the Eiffel Tower, she would likely have taken various aspects of her experience into account, such as the view from the top or the efficiency of the queueing process. However, the N-effect suggests that if Adriana were to post a review as a later contributor, her basis for evaluation would shift to focus more on her personal experience, rather than an analytical basis if she were an early reviewer.

Marketing strategies often aim to increase the number of opinions, based on the belief that aggregated, independent, and unbiased judgments are more accurate than individual ones - a phenomenon known as the wisdom of crowds (*Surowiecki, J. (2005). The Wisdom of Crowds. Anchor.*, 2005). However, it is unknown if opinions remain independent and unbiased as they accumulate. In large volumes, the possible influence of mass accumulation may go unnoticed, but this could matter if marketers and consumers infer conclusions and decisions from aggregated evaluations.

The N-effect is not just due to direct social influence. Instead, N serves as a cue for what contributors may focus on. As more opinions are added, contributors tend to focus more on their personal experience rather than analytically evaluating a target. This suggests that N plays a role in shaping the content of online opinions, beyond just social influence. Importantly, even if online contributors don't actually read any prior opinions, simply knowing that their N is low or high should affect their way of evaluating the target.

This work contributes to the literature on electronic word-of-mouth (eWOM) and online opinion sharing, particularly the role of group and audience size (Brewer & Kramer, 1986; Clarkson et al., 2013). Previous research has shown that consumers perceive a larger number of opinions as equal to more information, and N serves as a cue for crowd sentiment (Watson et al., 2018). However, I demonstrate that this assumption has limits. Past literature should consider this contextual factor, as a larger number of reviews does not necessarily imply independent opinions.

This work also contributes to the literature on consumer evaluations, which has shown that people rely more on analytical or emotional information depending on various factors, such as the type of product/service being reviewed (Kovács et al., 2014; Rocklage et al., 2018; Rocklage & Fazio, 2015) or individual differences in reviewers (Moe & Schweidel, 2012; Naylor et al., 2011). However, research has not considered the effect of N as a cue to contributors, an important factor as N continues to accumulate.

Theoretical Background

N in online opinions

In the next section, I will examine the role of N in shaping the type of evaluations that contributors provide on online opinion platforms. I will also explore how N influences the choice between analytical and emotional processing, and how this affects the content of contributors' evaluations.

In general, collective and unbiased judgments of a group of individuals are more accurate than those of an individual person (Makridakis & Winkler, 1983). This is because the average judgment from a group minimizes the biases and noise generated by individuals. This finding holds true in a wide range of situations—e.g., in sequential decision making, including expert probabilities of future events (Budescu & Chen, 2015), consumer choices (Salganik et al., 2006), competitive environments (Lichtendahl et al., 2013) , and crowdfunding (Polzin et al., 2018; Van de Rijt et al., 2014). The collection and interpretation of consumer online opinions may behave similarly. When audiences process online opinions, they use them, along with other cues, to arrive at a conclusion about overall sentiment based on the information provided.

Because greater N is perceived to be a signal of higher information quality, online platforms, like Yelp! and Google, display N prominently on their sites. Doing so increases the likelihood that N will be viewed by readers and contributors of opinions. Approximately 76% of consumers read reviews before making a purchase decision (Statista, 2021), and N has become a strong cue for product quality (de Langhe et al., 2016; Fisher et al., 2018; Watson et al., 2018). Awareness of N may begin at an early stage in the consumer journey due to the significant impact that online reviews have on the consumer decision process (Chevalier & Mayzlin, 2006; Chintagunta et al., 2010; Moe & Trusov, 2011).

I contend that a consumer's interpretation of N as few or many is influenced by the context. For example, a café with 200 reviews in a small town or village would be interpreted as having many reviews, while a café in central London with 200 reviews would be seen as having relatively few. The interpretation of N can also vary based on individuals' awareness of the size of the pool from which they are being sampled. For example, phrases such as "nationwide survey" and "local survey" can elicit different responses to the same questionnaire. Thus, I contend that the numeric value of N (e.g. 7, 700, or in some cases 7000) affects judgments and decisions through its "gist" interpretation, such as "very few" or "a great deal" (Reyna, 2009).

Moreover, N may change quickly and would not be the same for every contributor. It is not the total volume of opinions as seen by other readers, nor is it the passage of time, although time and N are positively correlated. Other interpretations of N could include the number of arguments made in favour or against a product, popularity, or the (un)certainty of quality. While there is merit in thinking about N in these ways, the present studies focus on the gradual build-up of consumer evaluations and specifically how this build-up affects the judgments and informational content that reviewers post.

Analytical vs. Emotional

Why are some evaluations more analytical and others more emotional? A key distinction in how opinions are generated is between experiential and rational processing when judging past experiences (CEST; Epstein et al., 1996; Epstein, 2003). Experiential processing, which is more affect-driven and rapid, involves evaluating past experiences based on feelings and emotions. On the other hand, rational processing is more analytical and effortful, requiring justification through logic and evidence. These two evaluation processes often work in synchrony, resulting in compromises between them (Epstein, 2003). In the context of online reviews, analytical and emotional thinking styles can be thought of as a continuum in two polar directions. Analytical thinking, for example has been shown to affect individuals' willingness to pay (Godek & Murray, 2008) and to help explain consumer attitudes towards the adoption of new technologies (Simon & Usunier, 2007). Differences in thinking styles can help us understand how contributors make their evaluations.

The distinction between analytical and emotional reviews is important because the way in which people write determines the influence of their message. Some research has highlighted the importance of detailed analytical information in fostering persuasion. Research has found that detailed information is more persuasive than general evaluative word-of-mouth (Feldman & Lynch, 1988a; Herr et al., 1991a; Lynch Jr et al., 1988). However, for experiential products, such as hotels, consumers tend to focus more on reviewer agreement rather than detailed information (Jiménez & Mendoza, 2013). Additionally, negative statistical reviews (such as numeric ratings) are perceived to be more credible than negative narrative reviews (Hong & Park, 2012).

On the other hand, other research has identified the role of emotions in facilitating persuasion (Petty et al., 1986; Wegener et al., 1994). In particular, work on the elaboration likelihood model (ELM) has discussed how both positive and negative emotions can enhance persuasion and the circumstances in which this occurs. This and other work has also drawn on an adaptive basis for emotions in persuasion. Research has suggested that outward displays of emotions evolved to influence others (Frijda & Mesquita, 1994) and individuals may use others' emotional expressions to inform their own attitudes (Van Kleef et al., 2015).

In sum, the way in which people write their online opinions and the type of information they provide can influence the effectiveness of their message. An analytical approach presents facts and logical arguments, while an emotional approach appeals to feelings. Consumers posting an online opinion should consider their audience and purpose when deciding which approach to use. For example, an analytical approach may be more suitable for a technical audience, while an emotional approach may be better for persuading an audience to act. It is important to consider both analytical and emotional reviews when evaluating the influence of online opinions. Table 1 further distinguishes these two writing styles.

Table 1: Comparison of the Emotional and Analytical Systems

Emotional	**Analytical**
Emotional; pleasure-pain oriented	Logical; reason oriented
Outcome oriented	Process oriented
Behaviour mediated by vibes from past experience	Behaviour mediated by conscious appraisal of events
More rapid processing; oriented toward immediate action	Slower processing; oriented toward delayed action
More crudely differentiated; broad generalization gradient; categorical thinking	More highly differentiated; dimensional thinking
More crudely integrated; dissociative, organized in part by emotional complexes (cognitive- affective modules)	More highly integrated
Self-evidently valid: "Seeing is believing."	Requires justification via logic and evidence

Note. Adapted from Cognitive-experiential self-theory: An integrative theory of personality by S. Epstein, 1991, in R. C. Curtis, editor, The relational self: Theoretical convergences in psychoanalysis and social psychology, New York: Guilford. Adapted by permission.

To capture these thought processes and evaluation styles, I use text-mining techniques and language style analysis (Hutto & Gilbert, 2014; Pennebaker et al., 2014) to distinguish between analytical and experiential evaluations. This approach allows me to differentiate between analytical and experiential evaluations, and provides a deeper understanding of the ways in which people are influenced by others and how they, in turn, influence others.

The N-effect

In the next section, I define N-effect. I also investigate drivers and moderators of the N-effect, as well as the role of contributor motivations and shared responsibility. I also outline hypotheses related to the N-effect and its impact on the formation and expression of opinions online. The aim is to deepen our understanding of this phenomenon and its relevance for online opinion platforms. First, I define the N-effect phenomenon.

When N is low, contributors are more likely to rely on analytical information and provide formal, logical evaluations. On the other hand, when N is high, contributors are more likely to rely on emotional information and provide personal, experiential evaluations. I dub this the N-effect.

For example, when asked to review their stay at a recently opened hotel with an online presence, low-N contributors are more likely to provide analytical evaluations that are well reasoned and that could include examples in support of their opinion. These evaluations may be useful for identifying specific areas for improvement and allowing consumers with different preferences to compare evaluations of the same hotel or across hotels. On the other hand, high-N contributors are more likely to provide experiential evaluations that focus on how the hotel made them feel and share their opinion through a narrative format. These evaluations may be more helpful for targets that elicit strong emotional responses or where the criteria for evaluation are unclear, such as an art exhibition or music events.

Despite completing the same task - providing an online opinion - the two evaluations produce very different content. As such I predict:

H1a: Contributors write more analytical content in online opinions when N is low in comparison to when N is high, for a specific target.

H1b: Contributors write more emotional online opinions when N is high in comparison to when N is low, for a specific target.

I will now delve into the underlying mechanism behind the prediction that N affects contributors' opinions, including the motivations of contributors to provide helpful evaluations, as past research suggests that a concern for others and the desire to have a social benefit are common motivations for online sharing (Cheung & Lee, 2012; Hennig-Thurau et al., 2004; Yoo & Gretzel, 2008).

How might the volume of opinions affect contributor evaluations? I argue when N is low, contributors' sense of responsibility to provide evaluative thoughts is strong and analytical accounts become the focus of their evaluation. N is processed heuristically and interpreted as a cue of the collective contribution of peers with similar motivations. Since analytical evaluations are more cognitively demanding and require greater effort than evaluations based on emotional information (Petty et al., 1980; Weldon & Gargano, 1985a), high-N contributors may focus on the latter while low N contributors may feel more responsibility and exert greater effort to provide analytical opinions.

Evidence supports the idea that shared responsibility leads to fewer evaluations and less complex judgment strategies. In experiments by Petty et al. (1980), participants who believed that many others were also evaluating a performance produced fewer evaluative thoughts than those who worked alone. Weldon and Gargano (1985) found similar results in a multi-attribute judgment task where awareness of shared responsibility led to fewer evaluations and less complex judgment strategies.

Online requests for help also demonstrate this trend. When contacted individually, people responded significantly more to an email request for help compared to those contacted in group emails (Barron & Yechiam, 2002). In virtual gaming environments, the presence of others decreased helpfulness (Kozlov & Johansen, 2010), and in live chat rooms, the number of participants correlated with the time it took to receive a response to a request for information (Markey, 2000). Therefore,

H2: Low-N contributors will feel greater responsibility to provide evaluative thoughts which are well-reasoned and logical in comparison to high-N contributors for a specific target.

In conclusion, as N increases, contributors' sense of responsibility to provide a well-reasoned and effortful evaluation diminishes, and they rely more on emotional experiences and provide less effortful evaluations. This highlights the importance of considering N when interpreting online reviews and the potential for biased evaluations.

In this section, I will examine ways to attenuate the difference in contributor evaluations caused by N. First, I will investigate the moderating role of social norms on the N-effect. Then, I will examine the impact of group objective on the sense of responsibility in online group contribution settings, focusing on the distinction between feedback and opinions. Finally, I will explore the influence of evaluating a target on multiple dimensions versus one dimension on the content of contributor evaluations.

Social norm as a moderator

The N-effect suggests an individual's sense of responsibility is influenced by the volume of opinions that appeared before their contribution. In this section, I will explore how social norms can moderate this effect and how providing clear expectations for evaluations can reduce the tendency to provide more emotional evaluations as N increases.

Social norms play a significant role in shaping our sense of responsibility. Research has demonstrated that social norms can influence prosocial behaviour (Berkowitz, 1972; Darley & Latane, 1968) and even alter our actions (Griskevicius et al., 2006). Consumer's provision of analytical evaluative opinions is heightened in low-N contexts, where a greater sense of responsibility exists. This is strengthened if the norm is emphasized, aligning with the consumer's own tendencies. However, as the number of opinions increases, this weakens, and the norm conflicts with the consumer's individual writing style, resulting in the disappearance of the effect on opinions.

Accordingly, I propose that the sense of responsibility is driven by social norms, and contributors interpret N as a cue for these norms. Reminding contributors of these norms in low-N scenarios can focus their attention on their sense of responsibility, while in high-N scenarios, the same reminder will have no effect due to the perceived change in social norms (Cialdini et al., 1990). When N is high, contributors are unlikely to change their opinion because they perceive the social norms as having a low sense of responsibility.

H3: Contributors' evaluation style is shaped by their perception of the norm, based on N. Low-N contributors perceive a norm for analytical evaluations and avoid violating it, while high-N contributors perceive a norm for emotional evaluations.

Moderating role of feedback vs. opinion prompt

Specific objectives are important in determining responsibility in online group settings. Previous research has shown that group size can decrease feelings of responsibility (Latané & Ni1981), but increasing group cohesiveness can reverse this effect. This may be because social norms create pressure to conform, and this pressure is stronger in larger groups (Rutkowski et al., 1983). Therefore, when people are part of a group with specific objectives, they are more likely to feel a greater sense of responsibility to meet those objectives.

It is important to distinguish between *opinions* and *feedback* in the context of online evaluations, or eWOM (electronic word-of-mouth). The Oxford dictionary defines an *opinion* as a subjective view or judgment, and *feedback* as objective information used to make improvements. In other words, *opinions* reflect a person's personal perspective, while *feedback* is more focused on providing specific, actionable information that can be used to improve products or services. This focus on objectivity and improvement can lead consumers to naturally refrain or reduce the number of emotional evaluative thoughts when providing feedback. For example, "This restaurant's pizza is terrible, the worst I've ever had." is an opinion, while "The pizza was overcooked and the crust was not consistent, it could use some serious improvement." is feedback. While online evaluations are often referred to as opinions, belonging to a group of either opinions or feedback can strengthen people's sense of responsibility to meet social norms. Based on this, *opinions* and *feedback* can be thought of as two distinct task objective. This means that people who are part of a group focused on providing feedback, rather than just expressing opinions, will feel a greater sense of responsibility to contribute useful, objective information.

Marketers and platform managers may seek to make improvements based on consumer views. In these cases, constructive feedback is perceived as more useful than just opinions (Oxford dictionary). Consumer online opinion sharing can provide feedback instead of just expressing opinions, helping companies make informed decisions about how to move forward.

H4: Contributors asked for "feedback" as opposed to "opinions" will provide more analytical content, thereby mitigating the impact of N.

Moderating role of rating scale format

The way a rating system is presented on a webpage can have a significant impact on consumers' evaluations, as they are often required to provide both a rating and a written review on the same page (Chen et al., 2018a; Schneider et al., 2021). In this section, I will explore how using a multidimensional rating system can affect the information that reviewers provide in their evaluations. For instance, Schneider et al. (2021) found that when participants were asked to rate a hotel on multiple dimensions such as cleanliness, comfort, and location, they provided more detailed and specific information about their stay. Similarly, (Chen et al., 2018a) found that using a multidimensional rating system led to more accurate and detailed ratings of a restaurant.

This idea is supported by the accessibility–diagnosticity framework, which suggests that environmental cues can influence how people judge a target and direct their focus to specific features (Feldman & Lynch, 1988b). More specifically, this framework suggest that responses can: firstly, be modified by the elicitation context and secondly, influence the inputs to judgment and the process acting on the those inputs (Feldman & Lynch, 1988b). For example, seeing a sale sign might make a consumer more likely to consider price as an important factor in their evaluation. The accessibility-diagnosticity model has been demonstrated to have a significant impact on judgments in surveys (Menon et al., 1995) and on WOM product judgments (Herr et al., 1991b). I therefore propose:

H5: A multidimensional rating structure will increase analytical content in online opinions when N is high.

In other words, requiring reviewers to rate a target on multiple dimensions versus one dimension activates cognition related to specific dimensions of the target (see Appendix E). As a result, reviewers are more likely to provide analytical content related to these dimensions. Making certain attributes of a target more accessible can influence the information that reviewers decide to share in their evaluations

Practical implications

My objective is to first explore the N-effect and its moderators. By understanding the moderators of the N-effect, I can identify specific conditions under which the N-effect occurs and tailor interventions accordingly. Additionally, I can gain insight into how factors such as the conditions or the context in which the review is being written may influence the relationship between the number of reviews and individual contributions.

For example, previous research has mainly focused on the reader's perspective and online behaviour (Chevalier & Mayzlin, 2006; Chintagunta et al., 2010; Duan et al., 2008; Li & Hitt, 2008). However, a recent meta-analysis found that review content significantly influences perceived helpfulness (Hong et al., 2017) and text attributes such as sentiment and readability can impact perceived helpfulness (Agnihotri & Bhattacharya, 2016).

Marketers should take note of the content of online opinions as it can influence consumer behaviour. Research suggests that emotional content can be more persuasive, while analytical content can make it easier to compare options. Marketers should adjust their strategies accordingly, for example by emphasizing emotional benefits if opinions are mainly emotional or product features if opinions are mainly analytical. The number of online opinions can also indicate the importance or interestingness of a news article (Tsagkias et al., 2010, 2011) and can significantly impact a company's advertising value (Korgaonkar & Wolin, 2002). This highlights the importance of understanding how the volume of opinions affects contributors, which I will address further in the general discussion.

Overview of Studies

In a series of 6 studies, I tested my predictions involving four stimulus domains. Study 1A and 1B examined the role of N on online opinions across restaurants and hotels using reviews from online platforms. I used text mining techniques to measure the level of emotionality and analytical thinking in each opinion. The objective was to provide evidence in a real-world context for H1a and H1b. In study 2, I aimed to improve the precision of the dependent measure and to test the proposed mechanism of N-effect, the sense of responsibility. Participants played a visual crossword game and their ordinal position was manipulated to examine the causal role of N on contributors' evaluations. Participants were asked to indicate how much of their opinion they were willing to dedicate to analytical and emotional content and how they preferred to compose their review.

In study 3, I examined how social norms moderate the relationship between N and contributors' sense of responsibility. Participants were asked to indicate how responsible they felt when writing a review for a product or service. Here I investigate whether social norms intensifies the relationship between N and contributors' sense of responsibility.

In studies 4 and 5, I examined practical moderators of the effect. In each, participants played a visual crossword game and experienced a virtual tour, respectively. Study 4 investigated how social cues moderate the relationship between N and contributors' sense of responsibility. The objective was to investigate how opinion collection purpose can influence the relationship between N and contributors' sense of responsibility. Finally, in study 5, I investigated the effect of rating structures on N-effect. Participants evaluated a virtual tour using either a single-dimensional rating scale or a multi-dimensional rating scale. The objective was to investigate whether a multi-dimensional rating structure can increase the amount of analytical content in opinions under high-N.

Studies 1A and 1B

In studies 1A and 1B, I examine the role of N on online opinions across restaurant and hotels sectors. Given restaurants and hotels are the most read online review categories[4], study 1A and 1B examine the influence of the accumulation of past number of evaluations for a given target on future evaluations for the same target. In study 1A, I analyse reviews form Yelp!.

[4] https://websitebuilder.org/blog/online-review-statistics/

Yelp is one of the most popular service review websites, with over 50 million unique users. It is one of the most well-known online review platforms with over 100 million reviews, 17% of which are restaurant reviews alone. Yelp, created in 2004, is an important source of e-WOM information (Babić Rosario et al., 2016; Zhou & Guo, 2017). Therefore, Yelp is an ideal setting to observe how the accumulation of evaluations can influence future consumer evaluations.

To provide further evidence of this effect in real world context, I collected further review data in hospitality in study 1B, using TripAdvior data. This secondary data set provides evidence that N-effect is not restricted to one consumer domain but is a widely applicable phenomenon.

In these studies, through the use of linguistic and sentiment text analysis technique on real online consumer online reviews, I predict that as the number of reviews accumulate for a given restaurant or hotel, later reviews will be more emotional and less analytical (H1). In order to capture sentiment, I employ text mining to measure the level in each written opinion posted by users (Ludwig et al., 2013; Rocklage et al., 2018). Focusing on the written opinion text of each reviewer allows me to measure variance between users' evaluation which cannot be done by analysing the numeric rating alone.

Data sets

Using a software program, I obtained 11,516 reviews across 40 restaurants for study 1A and, 9,118 reviews on 35 different hotels in study 1B. Reviews from different cities in the United States were chosen to enhance generalisability. For each review, I extracted variables related to the target under review and review characteristics (and additional reviewer characteristics for study 1A), allowing for control over a wide range of exogenous factors. Appendix A shows a sample review illustrating the variables extracted.

Determining Analytical

The "*Analytic*" measure in "LIWC" (Linguistic Inquiry and Word Count) is used to measure the analytical content of texts (Pennebaker et al., 2014). It captures the extent to which people use formal, logical, and hierarchical thinking patterns in their writing. Text that scores low in analytical thinking is written in more narrative ways. The *Analytic* measure is based on the Categorical-Dynamic

Index (CDI), which combines abstract and cognitive complexity. The output is a percentage, which reflects the degree to which the text expresses analytical thinking. It is important to note that the context or objective of the writing task can influence Analytic scores. For example, a persuasive essay may score higher in analytical thinking, while a personal narrative may score lower, see Appendix A. Nonetheless, it is possible to score analytic on a relative basis across texts generated from the same task.

Determining Emotionality

In order to measure the level of emotional content, I use VADER (Valenced Aware Dictionary and Sentiment Reasoner) (Hutto & Gilbert, 2014) sentiment analysis programming on Python to produce measures of valenced emotionality. VADER lexicon is a rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. VADER differentiates between positive and negative text but also indicates the strength of positive or negative sentiments, *Compound*[5]. *Compound* produces an output between -1(most extreme negative) and +1 (most extreme positive). For example, the sentence *"The food is really GOOD!"* has a *Compound*: 0.64, PosIndex: 0.513, NegIndex: 0.0, Neutral Index: 0.487 (See Appendix for table of examples). This can be interpreted as: 51.3% of the sentence is positive, 0% negative and 48.7% as neutral. *Compound's* score of 0.64 is an accurate predictor of the sentiment; it is able to capture the true emotional expression conveyed by the "!" and "GOOD" in capital letters. Using *Compound* as the dependent variable, I can determine if online evaluations gradually became more emotionally expressive as the number of reviews increases. However, because LIWC is more commonly applied in sentiment analysis research, I have included additional analyses using LIWC's *Tone* variable as a robustness check, see Appendix A.

Furthermore, for the measure of emotionality the data were segmented into valenced sentiment of consumers based on their star rating. Participants that gave ratings of either 1 or 2 stars were classified as "negative valenced reviewers". Participants that gave a star rating of 4 or 5 stars were classified as "positive valence reviewers" (see dummy description in next section). Lastly participants with a 3-star rating were excluded from the analysis because of the ambiguity of valence that follows

[5] It produces four sentiment scores: *Compound*, PosIndex, NegIndex and Neutral Index.

with 3-star reviews. These amounted to 1412 observation. I do not believe this significantly impacts the finding of this study.

Operationalisation of N

Following the logic set out in the previous section, I use *LogNthReview* which is the logged sequential position of each opinion for a particular target e.g., the 56th reviewer is position 56 of a given restaurant. I logged this variable to account for the high values of N across the sample of restaurants and hotels. Although I do not expect a significant difference between reviewers 300 and 320, I do expect a difference between reviewers 1 and 20. In other words, the effect of N on opinion content is non-linear.

Control Variables

To isolate the effects of cumulative number of reviews and the presence of N-effect, I controlled for review- and reviewer-specific variables as well as target specific fixed effects.

Review-specific controls.

Dummy Positivity is a dummy variable for ratings categorized as positively valenced. It takes value 1 for all those reviews that received either a 4 or 5 star rating. It takes a value of 0 for all those reviews that received either a value of 1 or 2 stars. Reviews that received a rating of 3-stars were left out in model 2 (outline below), as previously explained.

Other review-specific controls were *Elapsed days*$_k$, calculated as the number of days elapsed between the first postings and review "*i"* for the same restaurant. I include a time variable to account for any changes in technology, cultural trends and familiarity users gain with the use of online review platforms that may have arisen over the time span. *Word Count*$_k$, the number of words used in each individual review. I included this variable to control for the possibility that later reviews are longer/shorter in content, which allows for greater proportion of positive vs. neutral vs. negative content. Furthermore, short reviews are not a concern given the platforms' minimum character length requirement with each review. Nonetheless for robustness *Word Count*$_k$ was included.

Reviewer-specific controls.

In order to address potential selection factors in the types of reviewers that post reviews later (vs. earlier), I also controlled for a number of reviewer characteristics. Unfortunately the hotel dataset does not provide these controls. I believe this a limitation is overcome with the restaurant dataset which does contain multiple reviewer characteristics: *Reviews*$_i$ is the total number of reviews the user has completed on Yelp across time for any restaurant. *Friends*$_i$, is the number of "friends" each reviewer has on the platform. This controls for popular reviewers with followers on Yelp, where popular reviewers could influence other consumers before writing their own review. It is important to note that other control variables were left out of the analysis due to their lack of statistical significance.

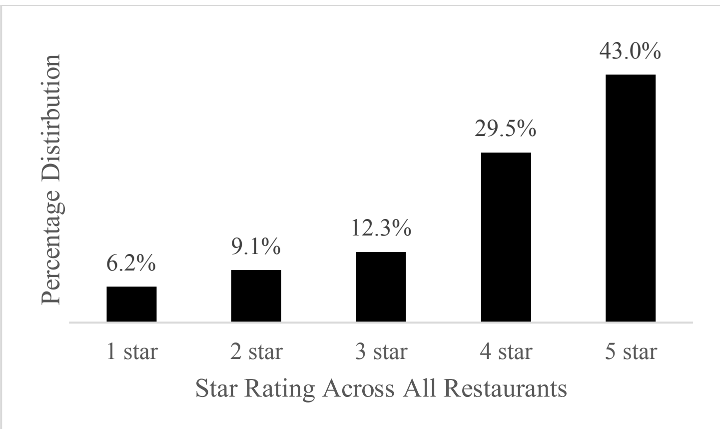
Target-specific controls. Restaurant/Hotel Average$_j$ is average rating for either restaurant or hotel j on a 5-point scale.

To address multicollinearity I opted for two methods: The first was analysing the correlation between control variables. No large correlation was found between the variables mentioned above. The second method was a study of the variance inflation factors. The results showed to be marginally above 1, indicating no multicollinearity (see table in Appendix A).

Study 1A: Evidence from Yelp restaurant evaluations

The average global restaurant rating is 3.94 out of 5 stars. The disproportionate number of 4 and 5 star reviews in our sample is consistent with past finding in other online platforms (He & Bond, 2015). Figure 1 below depicts the distribution of the five possible rating categories in the complete data set, including 3 star ratings.

Figure 1 . Overall Distribution of stars in Yelp dataset



Specifications

As an initial analysis the correlation between *(log)Nth review* and $Compound_{ijk}$ was 0.1 ($p< .01$) for positive reviews (i.e. 4 or 5 star rating), and -0.08 ($p< .01$) for negative reviews (i.e. 1 or 2 star rating). Recall that for negative reviews, extreme negative opinions tend toward -1. Opinions in the higher N range indicate a much stronger and faster move towards an extreme $Compound_{ijk}$ score. The correlation between *(log)Nth review* and $Analytic_{ijk}$ was -0.07 ($p< .01$) for positive and negative reviews. The low correlations may be due to a non-linear relationship between N and *analytic/compound*, as well as the inherent noise in the dependent variable. Additionally, other factors may play a larger role in this context, such as the high correlation between compound and star rating. These results are a model-free indication of an effect but further analysis is required to demonstrate the positive link between increasing number of reviews and consumer sentiment expression. The specification below describes a more rigorous approach.

The model specification includes $LogNthReview_i$ as our main independent measure. (See Appendix A for an additional specification using $NthReview_i$ for the robustness check). As previously mentioned I specify the two dependent measures, $Compound_{ijk}$ and $Analytic_{ijk:}$

$$Analytic_{ijk} = \alpha + \beta_1 LogNthReview_i + \beta_2 DummyPositiveReviews_k + \beta_3 LogInteraction\ term_{ik} + \beta_4 WordCount_k + \beta_5 Elapseddays_k + \beta_6 Restaurant\ Average_j + \beta_7 Friends_i + \beta_8 Reviews_i + \varepsilon_{ijk} \qquad (1)$$

$$Compound_{ijk} = \alpha + \beta_1 LogNthReview_i + \beta_2 DummyPositiveReviews_k + \beta_3 LogInteraction\ term_{ik} + \beta_4 WordCount_k + \beta_5 Elapseddays_k + \beta_6 Restaurant\ Average_j + \beta_7 Friends_i + \beta_8 Reviews_i + \varepsilon_{ijk} \qquad (2)$$

where "i" indexes the individual reviewer, j the restaurant under review, k the review specific effect, α is the constant and ε_{ij} the idiosyncratic error. In model (1) the dependent variable, is *Analytic* measured on a scale 0 to 100 where lower value suggest time-based stories and reflect a dynamic or narrative language style and higher value suggest formal, logical, abstract thinking and cognitive complexity (Pennebaker et al., 2014). Model (2) is the same as model (1) except $Compound_{ijk}$ ranges in a score between -1 and 1, where a score of 1 is considered to be a perfect positive opinion. Conversely, a score of -1 is considered to be a perfect negative expression. And 0 a perfectly neutral opinion.

Results

Results for the empirical models are presented in table 2 below. Results in table 2 support H1, the effect of the past number of reviews on online opinions, a strong negative relationship between Log*NthReview* and *Information* (β_1=-2.09, $p < .00$); with *Dummy Positivity* (β_2= 1.48, p=.57), *Log Interaction Term* (β_3=0.43, p=.44)*Word Count* (β_4= -.03, $p <.00$), *Elapsed days* (β_5=.00, p =.62) is statistically insignificant within the regression model, implying that the time between reviews has no effect on how consumers express themselves. The other control variables, *Restaurant Average* (β_6= -2.52, $p < .00$), *Friends* (β_8=-.00, p =.16), Reviews (β_9=.01, $p < .00$) (see table 2). I discuss the result of *Word Count* from both models in more detail below.

Regarding the results from model (2), a strong positive relationship between Log*NthReview* and *Compound* (β_1=-.06, $p < .00$); with *Dummy Positivity* (β_2= .27, $p < .00$), *Log Interaction Term* (β_3=.08, $p < .00$), *Word Count* (β_4= .00, $p < .00$), *Elapsed days* (β_5=.00, $p < .63$) is statistically insignificant within the regression model, implying that the time between reviews has no effect on how consumers express themselves. The other control variables, *Restaurant Average* (β_6= .00, p =.52), *Friends* (β_7=.00, p =.38), Reviews (β_8=.00, $p < .00$) (see table 2).

Table 2: Results for models 1 and 2

	Dependent variable:	
	Analytic	Compound
	(1)	(2)
Log Nth Review	-2.089***	-0.060***
	(0.524)	(0.010)
Dummy positive reviews	1.478	0.272***
	(2.634)	(0.047)
Log Interaction Term	0.431	0.084***
	(0.559)	(0.010)
Word count	-0.027***	0.002***
	(0.006)	(0.0001)
Elapsed days	0.0001	0.00000
	(0.0003)	(0.00000)
Restaurant Average	-2.524***	0.007
	(0.707)	(0.010)
Friends	-0.002	-0.00002
	(0.001)	(0.00002)
Reviews	0.006***	-0.0002***
	(0.002)	(0.00003)
Constant	74.821***	0.266***
	(3.470)	(0.057)
Observations	11,516	10,104
Log Likelihood	-54,852.850	-4,599.804
Akaike Inf. Crit.	109,723.700	9,217.608
Note:	*p**p***p<0.01	

Discussion

The results suggest that as the number of reviews increases, online opinions become more emotional and contain less analytical content. Analysis of model (1) shows that the percentage of analytical content provided in each opinion decreases with N. Other factors, such as *RestAverage* and *Reviews*, also appear to have an impact on the results. The negative relationship between *RestAverage* and analytical content may be due to a skewed distribution of positive restaurant ratings. The small impact of *Reviews* on *Analytic* may be influenced by the writing style of active Yelp users or the length of reviews consumers write.

In model (2), holding constant selection factors for the types of reviewers that rate early versus late, the coefficients on *DummyPositiveReviews* and the *InteractionTerm* suggest that *Compound* in online opinions trends towards extremes with every added review in a sequence. This indicates that N leads reviewers to express greater emotionality in their opinion, regardless of the valence of their experience.

Further analysis of *Word Count* revealed a small positive effect on *Compound* and a negative effect in model (1). This may be because longer opinions are more likely to contain content that is not solely analytical. It's important to note that the majority of online opinions (73%) are positively valenced, while only 15% are negatively valenced, leading to a highly skewed dataset towards positivity. Therefore, an increase in Word Count may be capturing the effect mostly driven by positive valenced reviews.

Study1B- Evidence from hotel evaluations

Specifications

Model (3) and model (4) present estimation results that include *Log NthReview* and the control variables with 9,118 observations in model (3) and 7,598 in model (4) due to the exclusion of 3 star ratings (see Study1A). *Compound* was developed using the same methodology stated earlier, through VADER sentiment analysis software. *Analytical* was generated following the same criteria as in Study1A. Because the hotel data contained fewer variables I modified the specification to fit the data available. However, as initial evidence in support of our hypotheses, *Compound* for positively valenced reviews had a correlation of 0.02 ($p<0.01$) with N and a correlation of -0.26 ($p<0.01$) for negatively valenced reviews with N. Whilst *Analytical* had a correlation of -0.04 ($p<0.01$) with N.

$$Analytic_{ijk} = \alpha + \beta_1 LogNthReview_i + \beta_2 WordCount_i + \beta_3 HotelAverage_i + \varepsilon_i \quad (3)$$

$$Compound_{ijk} = \alpha + \beta_1 LogNthReview_i + \beta_2 DummyPositiveReviews_i + \beta_3 LogInteraction\ term_i + \beta_4 WordCount_i + \beta_5 HotelAverage_i + \varepsilon_i \quad (4)$$

Results

Results for the second set of empirical models are presented in table below. Results in table 3 support model (3), the effect of the past number of reviews on online opinions, a strong negative relationship between Log*NthReview* and *Information* (β_1=-0.85, $p < .00$); with *Word Count* (β_2= .04, $p <.00$) and *Hotel Average* (β_4= 3.68, $p < .00$). I discuss the result of *Word Count* from both models in more detail below.

Regarding the results from model (4), a strong positive relationship between Log*NthReview* and *Compound* (β_1=-.04, $p < .00$); with *Dummy Positivity* (β_2= .71, $p < .00$), *Word Count* (β_4= .00, $p < .00$) and *Log Interaction Term* (β_5=.05, $p < .00$). The other control variable *Hotel Average* (β_4= .09, $p =.00$).

Table 3: Results for models 3 and 4

	Dependent variable:	
	Analytic (3)	Compound (4)
Constant	55.599***	-0.355***
	(1.923)	(0.055)
Log Nth Review	-0.854***	-0.042***
	(0.253)	(0.009)
Dummy positive reviews		0.713***
		(0.051)
Word count	0.035***	0.001***
	(0.007)	(0.0001)
Hotel Average	3.676***	0.090***
	(0.353)	(0.006)
Log Interaction Term		0.046***
		(0.010)
Observations	9,118	7,598
Log Likelihood	-43,263.220	-3,635.327
Akaike Inf. Crit.	86,534.450	7,282.654
Note:	*p**p***p<0.01	

Discussion

The results of the experiment suggest that as the number of reviews increases, online opinions become more emotional and less informative. Model (3) supports the hypothesis that the percentage of analytical content provided within each opinion decreases with N. Model (4) also shows that the trend towards extreme emotions in online opinions holds true regardless of the valenced experience of others.

Controlling for the characteristics of reviewers who rate early versus late, and the coefficients on *DummyPositiveReviews* and the *InteractionTerm*, model (4) demonstrates that N leads individuals to express greater emotionality in their opinions. Model (3) also suggests that as N increases, the amount of analytical content contained in each opinion as a percentage decreases.

As in Study 1A, these results suggest that N plays a significant role in shaping the content of online opinions for hotels. As individuals contribute more opinions, they may become more influenced by their own emotions and less by the analytical aspects of the target.

Does N cause individuals to change the content of their opinion?

I analysed two datasets of restaurant and hotel reviews and found that later opinions tend to be more emotional and less analytical compared to earlier opinions, regardless of the opinions of others (Raafat et al., 2009). This trend was observed for both positive and negative reviews, although the majority of reviews in our sample were rated positively. It is possible that this skewed distribution suggests a bandwagon effect or herding behaviour (Raafat et al., 2009), but I don't believe this to be the most accurate explanation. Previous research has shown that reviewer similarity declines over a sequence of reviewers (Godes & Silva, 2012), which suggests that many consumers do not read others' opinions before writing their own.

To confirm this, I segmented positive and negative reviews using the emotionality measure called Compound. This showed that sentiment measures were not converging to one extreme, but rather negative reviews were converging to a negative pole and positive reviews were converging to a positive pole as the number of reviews increased. The change in emotion and analytical thinking between opinions in an ordinal series was small, as expected. This finding prompted further investigation between high and low levels of N which I directly test in a series of experimental studies.

Both datasets have good external validity, meaning the results are likely to be generalisable to other similar contexts. However, it is important to consider the potential for unobserved effects that may have influenced the findings. For example, restaurants and hotels may have improved their service quality based on popular consumer feedback, which could explain why positive opinions become more extreme. On the other hand, negative opinions may reflect a higher level of disappointment if expectations consistently increase but are not met. Additionally, I cannot control for the time elapsed between the consumer's experience and the date of their review, or the various reasons that may influence how people express themselves. To address these potential alternative explanations and explore downstream consequences, I used more precise choice measures to predict individual writing in Study 3.

Study 2: Manipulating "N" for a review of visual crosswords

Study 2 was conducted during the Covid-19 lockdowns, when people were turning to the internet for entertainment. Participants played a visual crossword game, which is similar to a traditional crossword puzzle but uses paintings instead of words. The purpose of this study was to explore three objectives. First, I sought to examine the causal role of N by manipulating participants' ordinal position when evaluating the crossword game. Second, I aimed to improve the precision of our dependent measure by directly asking participants how much of their opinion they were willing to dedicate to an analytical evaluation (e.g., none to mostly) in a first measure and in a second measure how they preferred to compose their review (e.g., more analytical or more emotionally expressive). Finally, I aimed to test our proposed mechanism by measuring participants' sense of responsibility and testing its mediating role. I predicted that those at low-N would feel a greater sense of responsibility, which would increase their motivation to write in a more analytical and less emotionally expressive manner.

Method

Participants. U.S.-based participants on Mturk (N=183, M_{age}=37.8 , SD=10.17 , 55.7% male) completed the study for a small payment (AsPredicted #68546). No participants were dropped. The study had a 2-cell design with "N" manipulated between-subjects.

Material and procedure. At the study's start, participants were introduced to the new form of crosswords where the traditional format is replaced with famous paintings (see Appendix B). A common theme on overlapping squares of these images allows participants to complete the crossword. Then, they were told that they would play this game and could leave a review on our website, which would help others decide whether to download this game. Participants then clicked on a link to go through the visual crossword, where they completed on average 2 levels of difficulty. The game was brief (M_{time}= 2min, *SD*=2min18sec) after which they returned to the survey. Next, I asked participants to consider leaving a review of the game. I manipulated ordinal position ("N") between-subjects by notifying participants that they would be the 7th or 1105th person to post a review (low-N vs. high-N). I told participants before asking for their opinion, they would be asked how they would like to compose their review.

Measures. Participants were asked to write a brief headline of a few words to start their review. Then, for our key dependent measures, participants were asked how much of their review they would be willing to dedicate to "informational aspects", which they responded to on an unmarked slider scale, measured on a 7 point scale (1="No informational aspects", 7="Mostly informational aspects"). In a second key measure, participants were asked to indicate how they would like to compose their review. They responded to this on a scale in which they allocated 100 points to "informational aspect" versus "personal experience" . Finally, I measured responsibility in two steps: First, participants were asked in general, how much responsibility do you think a 1st reviewer would feel when posting their opinion? They answered on a 5-point scale (1= "no responsibility", 5 ="all responsibility"). I did this in order to address idiosyncratic perceptions of responsibility. Immediately after participants were asked a 4-item (α=0.92) scale to rate on a 7-point scale (1="much less", 7="much more"), how responsible do you feel to provide (1) an informative review for (2) a carefully thought-out review, (3) helpful review, and (4) a detailed review(see Web Appendix B for complete text on the measures). After responding to these measures, participants were participants and informed that only randomly selected people would be asked to write a review text and that they were not selected. Participants were then paid.

Results and Discussion

Our key dependent variable willingness to dedicate their opinion to informational aspects was subjected to an ANOVA, with N condition as an independent factor. As expected participants indicated that they would like to share more facts in low-N (*M*=5.35, *SD*=1.15) vs. high-N (*M*= 4.62, *SD*=1.62; $F(1,181)=12.18$, $p < .001$). The measure on relative allocation to providing informational aspect (vs. personal experience) in the review was subjected to an ANOVA, with N condition as an independent factor. As expected, participants focused their review relatively more on providing information in the low-N (*M*=56.8, *SD*=17.7) vs. high-N condition (*M*=47.97, *SD*=21.69; $F(1,181)=9.01$, $p < .00$), see table X. Conversely, allocation to expressing emotions was lower in low-N (*M*=43.2, *SD*=17.7) versus high-N (*M*=52.03, *SD*=21.69; $F(1,181)=9.01$, $p < .00$). In our analysis of responsibility participants indicated greater responsibility in low-N (*M*=5.4, *SD*=1.1) versus high-N (*M*=4.28 , *SD*=1.61 ; $F(1, 181)=29.61$, $p < .00$). I tested the hypothesised mechanism in a mediation analysis (Hayes 2008) employing Hayes (2012) PROCESS macro with bootstrapped samples (5,000). The model predicts our key dependent measure, relative focus on informational vs emotional content (trade-off scale), with independent factor of N condition (low vs. high-N) and responsibility as the mediator[6]. This analysis showed indirect-only mediation (Zhao et al., 2010). That is, with responsibility as the mediator, the effect of N condition on providing informational content was no longer significant ($\beta=-4.544$; $p > .1$), whereas the indirect mediation effect through responsibility was ($\beta=-4.283$; $p < .00$, 95% CI: [-7.365, -1.53]).

Table 4: Summary Statistics

	Mean	St.Dev	Median	Min	Max
LowN					
Facts	5.35	1.15	5.58	1.41	7.00
Information	56.80	17.70	57.50	1.00	99.00
Emotion	43.20	17.70	42.50	1.00	99.00
Responsibility base	5.40	1.10	5.75	2.25	7.00
Responsibility	5.40	1.10	5.75	2.25	7.00
HighN					
Facts	4.62	1.62	5.01	1.06	7.00
Information	47.97	21.69	50.00	0.00	89.00
Emotion	52.03	21.69	50.00	11.00	100.00
Responsibility base	4.28	1.61	4.00	1.00	7.00
Responsibility	4.28	1.61	4.00	1.00	7.00

[6] A secondary mediation analysis was carried out where I control for Responsibility base. This yielded the same results.

Study 2 experimentally showed that reviewers' awareness of their ordinal position affects how they write their opinions. This finding is consistent with our hypothesis that reviewers at low-N positions would strive to include more facts in their review. I attribute this N effect to differences in responsibility. Specifically, our results indicated that reviewers at low-N people felt greater sense of responsibility, which mediated the effect of ordinal position on review composition.

Study 3: Role of social norms

This study explores the relationship between social norms and individuals' sense of responsibility when providing online reviews. Previous research has demonstrated that social norms can significantly influence our behaviour (Goldstein et al., 2008). Social norms play a role in shaping our sense of responsibility, as seen in research on prosocial behaviour and altered actions (Berkowitz, 1972; Darley & Latane, 1968; Griskevicius et al., 2006). I argued that consumers tend to provide more analytical opinions in low-N contexts, where a greater sense of responsibility exists and is strengthened by emphasizing the norm (H3). However, as the number of opinions increases, the norm conflicts with the consumer's writing style, resulting in the disappearance of the effect. I therefore propose that reminding contributors of social norms in low-N scenarios can focus their attention on their sense of responsibility. In high-N scenarios, the same reminder will have no effect due to the perceived change in social norms. Contributors are unlikely to change their opinion because they perceive the social norms as having a low sense of responsibility.

Method

Participants. U.S.-based participants on Prolific Academic (N=232, Mage= 36.4, *SD*= 8.04, 44% male) completed the study for a small payment (AsPredicted #38647). The study was pre-registered (see Web Appendix A). The initial sample was *N* = 242, but 10 participants were excluded according to our pre-set criteria for none-sense writing/copy past and attention check. All results hold when including these 10 participants in the analysis. The study had a 2-cell design with "N" manipulated between-subjects.

Material and procedure

At the study's start, participants were introduced to the phenomenon of venues (museums, galleries, etc.) promoting their virtual tours in order to maintain engagement with the public during the Covid-19 pandemic. Then, they were told that they would view a virtual tour and could leave a review on our website, which would help the museum design its online presence. Participants then clicked on a link to go through the virtual tour of a natural history museum in the U.S., where they navigated between rooms and saw various exhibits. The virtual tour was brief (M_{time}=3min, SD=2min45sec) after which they returned to the survey.

Next, I asked participants to consider leaving a review of the virtual tour. I manipulated ordinal position ("N") between-subjects by notifying participants that they would be the 5th or 750th person to post a review (low vs. high-N). I told participants they could either post their opinion directly on our website or simply write it in our survey and I would upload it for them. To enhance believability, participants also visited our website where they could confirm that their reviewer position was 5 or 750. I created two versions of the website with the appearance of more versus fewer past reviews. To eliminate any effect of others' opinions, participants could not read all the past reviews. Crucially, participants were then told that their review was analysed using a text analysis software "LIWC" to provide insights. They spent a few second on a page indicating "LIWC analysis in Progress…" before moving on. Low-N and high-N conditions were informed that their opinions were relatively more emotional than analytical.

Measures

Once told about the content of their opinion participants were asked if they would like to change what they had written. Participants responded using a trade-off scale where they were asked to indicate how they would like to compose their review. Specifically, they allocated 100 points to "providing information" versus "expressing emotion", on a two-item scale. In item 1, they split 100 points between "adding detailed description of the virtual tour" versus "adding emotional impact." In item 2, they split 100 points between "providing information" and "giving more of a sense of how you felt." I combine the two items for our analysis (α=0.72), where higher numbers indicate a relatively greater allocation towards providing information. In a separate measure, participants were asked how many facts they would include in their review, which they responded to on an unmarked slider scale, which was measured on 100 points (0="add none", 100="add a lot").

Results and Discussion

The measure on relative allocation to providing analytical in the review was subjected to an ANOVA, with N condition as an independent factor. As expected, participants focused their review relatively more on providing information in the low-N (M=58.71, SD=18.3) vs. high-N condition (M=52.38, SD=20.36; $F(1, 230)=5.86$, $p< .02$), see table 2. Conversely, allocation to expressing emotions was lower in low-N (M=41.28, SD=18.3) versus high-N (M=47.62, SD=20.36; $F(1, 230)=5.86$, $p<.02$). Participants also indicated that they would like to share more facts in low-N (M=49.14, SD=24) vs. high-N (M=35.3, SD=24; $F(1, 230)=18.67$, $p<.01$).

Our study investigates the relationship between social factors and individuals' sense of responsibility when providing online reviews. By examining N on the content and tone of individuals' reviews, we gain insights into how social norms shape individuals' behaviour in online settings. Study 3 finds that social norms play a crucial role in determining individuals' sense of responsibility in both low-N and high-N contexts.

In low-N scenarios, social norms exert a stronger influence on individuals, leading them to feel a greater sense of responsibility for the content of their reviews. This is reflected in the increased focus on providing analytical content and less on expressing emotions, as well as a desire to share more facts. This suggests that reminding individuals of social norms in low-N situations can focus their attention on their sense of responsibility. In contrast, when N is high, the perceived social norm changes and the influence of social norms on individuals is less pronounced.

Overall, our results contribute to a better understanding of how social norms influence individuals' sense of responsibility when providing online reviews. While study 3 provides insight into the mechanism driving the N-effect, it does not take address on practical factors that may influence N-effect. I look at these in the studies that follow.

Table 5: Summary Statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
LowN Condition						
Information	58.549	18.306	12	50	70	98
Emotionality	41.451	18.306	2	30	50	88
Facts	49.127	24.680	1.000	32.267	64.860	100.000
Credibility Concern	3.960	1.293	1.000	3.116	4.879	6.600
HighN Condition						
Information	52.379	20.375	2.500	37.250	69.625	96.000
Emotionality	47.621	20.375	4.000	30.375	62.750	97.500
Facts	35.297	24.065	1.000	15.152	51.130	100.000
Credibility Concern	2.804	1.313	1	1.8	3.8	6

Study 4: Moderating role of feedback vs. opinion

The purpose of this study is to investigate how the objective of opinion collection for a visual crossword game affects the evaluation content. Specifically, I manipulate the framing of reviews as either “feedback”, which is a critical assessment that is constructive and neutral, or “opinions”, which are individual beliefs or sentiments that are commonly seen in online reviews.

Online reviews provide valuable insights into the quality and value of products or services for consumers. However, the content of these reviews can vary greatly depending on the context in which they are written. In this experiment, I examine how the distinction between “feedback” and “opinions” can influence contributors. I predict that participants prompted to think of their evaluation as “feedback” will provide more analytical evaluations regardless of N, while participants prompted with an “opinion” objective will provide evaluations as predicted by the N-effect (H4). Previous research has suggested that the way reviews are framed can affect their content (Cialdini & Goldstein, 2004). For example, framing a review as “feedback” may lead to more constructive and actionable criticism, while framing it as an “opinion” may result in more subjective and emotive language.

Participants in this study followed the same procedure as in study 3, experiencing the same visual crossword game. In study 4 I compare the effect of two prompts on the N-effect: one for leaving a review for other consumers (opinions) and another for helping developers improve the game (feedback). These prompts represent common motivations for leaving reviews on different online platforms. I experimentally manipulated participants' ordinal position and presented the prompts before the main dependent measure (see Appendix D).

Method

Participants. U.S.-based participants on Prolific Academic (N=259, M_{age}= 38.7, *SD*= 10.1, 30.5% male) completed the study for a small payment (AsPredicted #69595). The study was pre-registered (see Appendix D). The initial sample was *N* = 278, but 19 participants were excluded according to our pre-set criteria for none-sense writing/copy past, attention check, study duration above 20min or under 3min. The study had a 2 (N: low-N vs. high-N) X 2 (framing of task: feedback versus neutral) between-subjects design.

Material and procedure. At the study's start, participants were introduced to the new concept of visual crosswords. Then, they were told that they would view the game and could leave a review. Participants then clicked on a link to go through the visual crossword where they saw multiple art works and attempted 2 levels. The crossword game was brief (M_{time}= 3min 13seconds, *SD*= 2min7sec) after which they returned to the survey.

Next, there was a between-subjects manipulation of task framing. Half the participants (*opinion prompt condition*) were told that their review would help others decide whether to download the game. The other half of participants (*feedback prompt condition)* were told that their review would help developers improve the gaming experience for future users by fixing bugs and to keep features that performed well, and that developers were making changes continuously as feedback was being provided. I manipulated ordinal position ("N") between-subjects by notifying participants that they would be the 7th or 1105th person to post a review (low vs. high-N).

Measures. Participants were asked to write a brief headline of a few words to start their review. Then, for the key dependent measures, participants were asked how many facts they would include in their review, which they responded to on an unmarked 7-point continuous slider scale (1="None", 7="Mostly"). As in study 3 in a separate measure, participants were asked to indicate how they would like to compose their review. They responded to this on a scale in which they allocated 100 points to “informational aspects” versus “personal experience”, where higher numbers indicate a relatively greater allocation towards providing information. Finally, participants were thanked and informed that only randomly selected people would be asked to write a review text and that they were not selected for this additional task. Participants were then paid.

Results and Discussion

The measure on relative allocation to providing information (vs. expressing emotion) in the review was subjected to an ANOVA, with N condition as an independent factor. As expected, for the opinion condition, participants focused their review relatively more on providing information in the low-N (M=55.65, SD=19.55) vs. high-N condition (M =44.55 , SD=20.30; $F(1,122)=9.62$, $p <.001$). In the feedback condition low-N (M=53.65, SD=20.8) vs. high-N condition (M =52.03 , SD=21.13; $F(1,133)=0.195$, $p<.66$) showed no significant difference. Furthermore, as predicted there was a weak moderation effect of the prompt type on N condition (b=9.47, SE=5.14, $p < .06$). A contrast analysis indicated that under high-N participants would share more analytical (M=52.03, SD=21.44) for the feedback condition in comparison to the opinion condition (M= 44.55 , SD=20.30; $F(3,255)=3.42$, $p <0.02$).

Lastly, when asked how many facts they would include in their opinion, participants in low-N (M=4.92, SD=1.25) versus high-N (M=4.21, SD=1.46; $F(1, 122)=8.59$, $p< .00$). Whereas in the feedback condition low-N (M=4.95, SD=1.25) versus high-N (M=4.67, SD=1.46; $F(1, 133)=8.1$, $p <.26$), no significant difference. This last results further support H4. However there was no moderation effect of the prompt type and N condition of facts (b=0.44, SE=0.35, $p <.20$).

This study aimed to clarify the effect of different objectives on consumer evaluations of a target. The results showed that under low-N the use of an *opinion* or *feedback* prompt does not significantly impact the content of contributors' opinion as predicted by N-effect. This is likely due to a high sense of responsibility among contributors to focus on sharing analytical evaluative thoughts. However, when N is high, the use of a *feedback* framing results in an increase in the analytical content of contributors' opinions. The results also indicated that using an *opinion* prompt under high-N does not produce distinct outcomes compared to a typical review request.

The results of this study provide several key conclusions. Firstly, the use of constructive feedback prompts encouraged participants to prioritize the provision of analytical evaluative thoughts, despite the increased cognitive effort required. Secondly, participants disregarded the social norm discussed in study 3 and opted for analytical thoughts, even if it was not in line with their personal preference. Finally, when asked directly about the number of facts to include, there was no difference between low-N and high-N conditions.

These findings highlight the impact of prompts in shaping online opinions and how they can provide readers with a more analytical-oriented evaluation. Despite their potential to greatly influence the writing process, the use of prompts in online review platforms is often limited and may go unnoticed by consumers. In the following study, I explore the effect of a more prevalent factor, rating structures, on the writing process and offer valuable insights for platform design.

Study 5: Single vs. Multi dimension ratings

Study 5 is designed to demonstrate the moderating role of rating structures—single versus multidimensional—on the N-effect. Ratings structures are relevant because online platforms obtain opinions through both ratings and opinions, which are often collected at the same time. Most platforms use a single-dimensional rating system, where consumers evaluate the target on a single 5-star scale. However, multi-dimensional scales have gained widespread popularity on many of the most prominent websites, e.g. TripAdvisor, Booking.com and OpenTable (see Appendix E). According to hypothesis H5, I argue that in a multidimensional rating structure, contributors' opinions tend to focus more on the specific aspects rated, as these informational aspects are more accessible at the time of evaluations (insert reference). This should attenuate the N-effect, leading to a similarly high focus on providing informational rather than emotional content, irrespective of N.

The study took place during the Covid-19 lockdowns, when museums were encouraging virtual visits. Participants experienced a virtual tour of a museum and then reviewed it. I manipulated reviewers' ordinal position and presented the rating structure as either single dimension (one 5-star scale) or five dimensions (five 5-star scales). Participants could post their review on our website, where I controlled their position and the N they believed they were in (low or high). (See Appendix E for further details).

Method

Participants. U.S.-based participants on Prolific Academic (N=420, M_{age}= 37.93, *SD*= 9.1, 48.8% male) completed the study for a small payment (AsPredicted #40042). The study was pre-registered (see Appendix A). The initial sample was 499N, but N = 77 participants were excluded according to our pre-set criteria for nonsense writing/copy past and attention check. The study had a 2 (N: low-N vs. high-N) X 2 (rating dimensions: 1-dimension versus 5-dimensions) between-subjects design.

Material and procedure. At the study's start, participants were introduced to the phenomenon of venues (museums, galleries, etc.) promoting their virtual tours in order to maintain engagement with the public during the Covid-19 pandemic. Then, they were told that they would view a virtual tour and could leave a review on our website, which would help the museum design its online presence. Participants then clicked on a link to go through the virtual tour of a natural history museum in the U.S., where they navigated between rooms and saw various exhibits. The virtual tour was brief (M_{time}= 3min, *SD*= 2min42sec) after which they returned to the survey.

Next, there was a between-subjects manipulation of rating structure. Half of the participants were asked to rate the tour on a 5-star scale (*single dimension rating structure)*. The other half of participants rated the tour on 5-point scales across five aspects: overall, image quality, ease of navigation, collection of exhibits and layout of virtual tour webpage (*multidimensional rating structure).* I manipulated ordinal position ("N") between-subjects by notifying participants that they would be the 7th or 690th person to post a review (low vs. high-N). I later asked participants to consider leaving a review and that they could either post their opinion directly on our website or simply write it in our survey and I would upload it for them. To enhance believability, participants also visited a website where they could confirm that their reviewer position was 7 or 690. I created two versions of the website with the appearance of more versus fewer past reviews. To eliminate any effect of others' opinions, participants could not read all the past reviews.

Measures. Participants were asked to write a brief headline of a few words to start their review. Then, for the key dependent measure, participants were asked to indicate how they would like to compose their review. They responded to this on a scale in which they allocated 100 points to "providing information" versus "expressing emotion" in two items. In item 1, they split 100 points between "adding detailed description of the virtual tour" versus "adding emotional impact." In item 2, they split 100 points between "providing information" and "giving more of a sense of how you felt." I combine the two items for the analysis (α=0.72), where higher numbers indicate a relatively greater allocation towards providing information. In a separate measure, participants were asked how many facts they would include in their review, which they responded to on an unmarked slider scale, which was measured on 100 points (0="add none", 100="add a lot"). After responding to these measures, I thanked participants and invited them to post a review of the virtual tour on our website if they like, or to finish the study for payment.

Results and Discussion

The measure on relative allocation to providing information (vs. expressing emotion) in the review was subjected to an ANOVA, with N condition as an independent factor. As expected, participants focused their review relatively more on providing analytical thoughts in the low-N (*M*=53.37, *SD*=20.18) vs. high-N condition (*M*=48.97, *SD*=18.65; $F(1,201)=3.54$, $p< .06$) in the single dimensional rating structure. Furthermore, the interaction between N conditions and ratings structure was also significant ($F(1, 416)=4.98$, $p <.03$). Lastly, a contrast analysis indicated that participants would share more analytical content in high-N for the multidimensional ratings structure (*M*=54.94, *SD*=21.35) (and less emotion (*M*=45.06, *SD*=21.37)) in comparison to high-N in the single dimension rating structure for analytical (*M*= 48.97 , *SD*=19.02; $F(1,206)=4.55$, $p< 0.03$) (and emotion (*M*= 51.03, *SD*= 19.02)).

Study 5 replicates earlier findings that reviewers consider their ordinal position (N) when composing their review (H1a and H1b). The findings also suggest that the number of idea ratings (in this case 1 versus 5) can have a significant effect on evaluations (Riedl et al., 2013). Specifically, the study finds that a multidimensional rating structure, in comparison to a single dimension, can increase the amount of analytical thoughts in contributor opinions under high N (H5). These results align with previous research showing that multidimensional scales can lead to more consistent results and are less susceptible to response bias (Archak et al., 2011; Chen et al., 2018b; Moe & Schweidel, 2012). Single-criterion evaluations can be ambiguous tasks, causing contributors to base their evaluations on cues such as individual dimensions of a multidimensional structure (Christian et al., 2007).

General Discussion

The studies presented here have examined the dynamics of online opinion expression in response to volume. I have presented a novel phenomenon referred to as the "N-effect" that demonstrates how online opinions are influenced by the accumulation of other opinions. By examining the impact of N on opinion content, I have shown that consumers provide less analytical and more emotional evaluations as N increases. This suggests that N is an imperfect signal of information reliability in online settings and that online platforms should not equate the volume of reviews with the overall consistency of the information collected.

Importantly, I have demonstrated that the N-effect is not just due to direct social influence, but also serves as a cue for what contributors may focus on. As more opinions are added, contributors tend to focus more on their personal experience rather than analytically evaluating a target. This has important implications for marketers and consumers as they infer conclusions and decisions from aggregated evaluations. This work contributes to the literature on eWOM and online opinion sharing by highlighting the limitations of using N as a signal of information reliability.

Results show that as the volume of online opinions increases, they become more emotional and less analytical. This trend was observed in both positive and negative reviews in two datasets of restaurant (study 1A) and hotel reviews (study1b). In order to establish a causal relationship and rule out alternative explanations for the N-effect, studies 1A and 1 B controlled for target-specific, reviewer-specific, and review-specific aspects. Study 2 finds that a contributor's sense of responsibility towards others decreases as N increases, with reviewers at low-N positions feeling greater sense of responsibility, focusing on analytical content and less on expressing emotions. Additionally, the use of a choice variable in study 2 improves the robustness of the results. Overall, these studies provide insights into the mechanisms driving the N-effect and contribute to a better understanding of how individuals' sense of responsibility influence online opinions. However, it's important to note that more research is needed to explore the practical factors that may influence N-effect.

I further explore how perceived social norms, prompts, and rating structures affect online reviews. Study 3 finds that social norms impact individuals' sense of responsibility in low and high-N scenarios. In low-N situations, social norms emphasize responsibility lead to more analytical evaluative thoughts and less emotional expression, while in high-N scenarios reviewers are less affected by this unaligned social norms. Study 4 demonstrates that prompts can shape opinion content; a prompt focused on feedback (vs. opinion) increases analytical content when N is high through an increase in sense of responsibility. Study 5 confirms that N plays a role in shaping the content of online reviews and suggests that multidimensional rating structures can increase the amount of analytical content included in reviews under high-N.

Finally, the use of two different stimuli (a crossword game and virtual tour) in the experimental studies (study 2, 3, 4 and 5) increases the generalizability of the findings. Furthermore, the use of both text and choice variables as dependent measures further strengthens the robustness of the findings. These results provide insights into the mechanisms driving the N-effect and contribute to a better understanding of how N influences online opinions.

Theoretical Contributions

This work makes several contributions. Theoretically, it adds to the literature on online social influence and online opinion sharing by examining how N is likely to influence contributor judgments as a signal of collective contributions by others. While previous research has interpreted the number of opinions as economic indicators of performance (Chevalier & Mayzlin, 2006; Dellarocas et al., 2007; Duan, Gu, & Whinston, 2008; Li & Hitt, 2008; Liu, 2006; Moe & Trusov, 2011; Zhu & Zhang, 2010), none to our knowledge have taken the perspective of the opinion contributor. The present studies confirm that the volume of evaluations, or accumulation of opinions, does influence contributor opinion content.

Secondly, this research shows that consumer provided opinions can be more emotional or analytical (Epstein, 1996). In the context of online evaluations, the current project adds to our understanding of how online opinions are contributed. Specifically, it highlights the importance of post-purchase written evaluations in shaping the opinions of potential readers. This is the first study to examine the impact of review accumulation on online opinion and to highlight the influence of post-purchase written evaluations on others.

Online opinion content can have a significant impact on perceptions and decision-making of potential customers. On one hand, previous research has shown that affective-rich content in online reviews can have a greater influence on customers (Ludwig et al., 2013; Rocklage & Fazio, 2018). However, it's important to note that excessive affective arousal can also lead to a negative effect on review helpfulness (Yin et al., 2017). On the other hand, a higher proportion of factual statements in online reviews has been found to aid readers in their decision-making, for example with books and cars (Schindler & Bickart, 2012). These conflicting findings highlight the need for a deeper understanding of the dynamics of online opinion expression and the role of their content.

Furthermore, this work provides preliminary evidence in an online context of how a reduced sense of responsibility may lead to affect-rich judgments. Importantly, it also outlines conditions where online groups may retain a strong sense of responsibility and thus produce analytical evaluations of targets despite a high number of contributions from others. This extends the group size literature by demonstrating how responsibility may be lost in online contexts when group members blend together (Baumeister et al., 2016). Previous research on online reviews has not considered the impact of such cues on contributors, but group size research suggests that individuals in larger groups may reduce effortful evaluations relative to small groups (Weldon and Gargano 1985).

One reason online platforms aggregate many opinions is driven by the belief that larger samples will result in a more accurate average opinion that reflects the true value of the target. Research has shown that when multiple individuals make a judgment, their average is often more accurate than most individual judgments (Surowiecki, 2004; Sunstein, 2006). This is known as the wisdom of crowds, where an individual can consult and combine the opinions of others to better estimate the true value of a target. However, the size of the group may also bias the contributions of group members and lead to conformity through phenomena such as bandwagon or herding behaviour (Milgram et al., 1969). My research provides a qualitative perspective on the wisdom of crowds in online settings by emphasizing the limitations of using N as a signal of information. It therefore may be useful to consider the wisdom of crowds from a qualitative perspective, such as through text analysis.

The findings on task objective (study 4) align with previous research in the field of visual design, which has shown that symbols and graphics can significantly impact how respondents answer surveys (Christian & Dillman, 2004; Dillman & Christian, 2005; Redline et al., 2003; Tourangeau et al., 2004). These results demonstrate that when multiple options for providing an opinion are available, survey designers have the ability to manipulate the instructions to obtain a specific format (Christian et al., 2007). In the context of online opinion expression, contributor evaluations can be ambiguous and the criteria for evaluation can vary between individuals. The use of prompts in this study highlights the importance of standardizing judgment criteria to a specific content type, while still allowing for the preservation of contributor uniqueness.

In addition, the findings on the impact of multidimensional rating structures (study 5) provide further insight into how these structures can shape online opinion expression. The use of multiple rating criteria may facilitate preference matching and uncertainty reduction, which allow consumers to process rating information more efficiently (Chen et al., 2018) . It is also possible that additional rating dimensions could simplify the evaluation process by making these dimension more accessible (Schneider et al., 2021). These findings offer practical implications for platform design and add to the current understanding of online behaviour.

Managerial Implications

Online review volumes have become a crucial tool for businesses and firms to attract consumers. While the valence of reviews can impact consumers early in their decision-making process, the volume of reviews has become a standalone indicator of trust in online reviews and a cue of product quality for consumers. As online review volumes continue to rise, it is increasingly important for businesses to have strategies in place to address such volumes and the impact they may have on contributors. The presentation of the appropriate evaluation content can significantly affect decision outcomes, as consumers typically read only between one and ten reviews (Statista 2021).

Marketers and online platforms should take into consideration the balance between analytical and emotional opinions in online reviews as they can result in different types of persuasive information. Past research has found that detailed information about a product is more persuasive (Bansal and Voyer 2000; Feldman and Lynch 1988; Herr, Kardes, and Kim 1991; Lynch, Marmorstein and Weigold 1988). On the other hand, other works have shown that emotions can enhance the persuasiveness of a message (Petty & Cacioppo, 1986; Wegener & Petty, 1994). While this study does not explore persuasiveness, it does provide methods for achieving the desired type of information in online reviews.

The value of a review depend on the type of product being reviewed (Mudambi & Schuff, 2010). For example, if a company is launching a new technology product, it may be more beneficial to highlight the detailed features and specifications through analytical opinions in order to persuade consumers that the product is reliable (Schindler & Bickart, 2012). However, if the company is promoting a new luxury experience, emotional opinions that highlight the brand's exclusivity and emotional appeal may be more effective in persuading consumers to purchase the product.

Further, analytical evaluations can be used to make measurable, observable improvements and allow consumers with different preferences to compare evaluations of the same hotel or across different hotels (Baek et al., 2012). Emotional evaluations may be more helpful in situations that trigger emotional responses or where criteria for evaluation are undefined and ambiguous, such as when evaluating a new restaurant or a local art gallery (Huang et al., 2013).

It is worth noting that product categories may not be the only factor influencing evaluations style. Over time, as review for a given target accumulate the utility of emotional reviews can decrease (Agnihotri & Bhattacharya, 2016). Analytical evaluations may help address this issue. For example, a recently renovated hotel, typically a hedonic product, may benefit from analytical content focusing on key attribute improvements, as it increases believability of quality improvements made by management. On the other hand, traditional utilitarian products, such as cooking tools, may benefit from positive emotional accounts from consumers, highlighting standout benefits.

This work provides practical insights for online platform managers on how to interpret and utilize consumer-written reviews and highlights the potential biases that can arise from interpreting these reviews. Managers' tendency to misinterpret or ignore warnings (Tinsley, Dillon, & Madsen, 2011) can be compounded by the potential for consumer-written expressions to lead to distorted perceptions (Carter et al., 2007; Franzosi, 1998; Gorham, 2006). Managers also face the challenge of accurately interpreting consumer-driven information, as the language style used by consumers can make feedback more complex, especially when emotions are involved (Antioco & Coussement, 2018).

Managers often rely on their intuitions to interpret consumer evaluations, but this can lack clear reasoning (Schoemaker & Russo, 1994). From a survey quality perspective, marketers can minimize the type of content that consumers choose to focus on in their responses, reducing survey error (Biemer, 2010). Here I show evidence of an unintended and undocumented (qualitative) measurement error—consumers' increasing reliance on emotional over analytical content to express an opinion. The findings in this work can be used by online platform managers to mitigate their own judgment biases when interpreting consumers' written reviews.

The findings of this research have practical implications for online platform managers as they interpret and utilize consumer-written reviews effectively. The potential biases that can arise from these reviews, such as misinterpretation or ignoring of warnings (Tinsley et al., 2011), and distorted perceptions (Carter et al., 2007; Franzosi, 1998), pose a challenge for managers. The language style used by consumers in their feedback, especially when cognitive thoughts are involved, can also make interpretation complex (Antioco & Coussement, 2018). Marketers thus often rely on their intuitions to interpret consumer evaluations, but this can lack clear reasoning (Schoemaker & Russo, 1994). This research presents evidence of an unintended and undocumented (qualitative) measurement error - the increasing reliance of consumers on emotional content over analytical content when expressing opinions. These findings can help online platform managers to mitigate their own judgment biases when interpreting consumers' written reviews.

This work also provides alternative guidelines to online platform managers on how to present the growing volume of reviews. Sorting reviews can significantly impact what consumers see – on average, 99.2% of businesses on Yelp are hidden when users sort results by "Highest Rated" or "Most Reviewed"[7]. My work emphasises the role of review content and the diagnostic value of opinions. Platforms should consider the volume of opinions for different product categories. For example, imagine an online platform for reviewing car dealerships. The platform manager may notice that many of the reviews are becoming increasingly emotional as the volume grows. To address this, the platform manager could implement a sorting feature that allows users to filter reviews based on their level of analytical content. This way, users looking for more in-depth and analytical reviews about a particular dealership would be able to easily find them.

Finally, review platforms may use N as a key performance indicator (KPI) to show consumers how well they are doing in accumulating opinions and engagement on the platform. The more opinions a platform has, the more active and engaged its users are likely to be, which can be attractive to potential consumers. However, research has not yet considered the impact that the accumulation of opinions may have on this metric.

[7] https://www.thinktur.org/media/onlinereview-infographic-1.pdf

Chapter 3: The Effect of User-generated Photos on Review Value

Caveat Spectator

Introduction

Many popular online review platforms allow users to post photos alongside their review text and rating. See Table 1 for an overview. Photos may depict locations, product sizes and materials, at-home usage, or other key aspects. Still, review platforms are unclear on effective policies for photos (Baymard Institute 2020)[8]. Of course, photos could reveal or illustrate visual information (Marder, Erz, Angell, and Plangger 2019; Xia, Pan, Zhou, and Zhang 2020), but their other impacts on review readers are not well understood.

Table 1

CHARACTERISTICS OF USER-GENERATED PHOTOS ON POPULAR REVIEW PLATFORMS

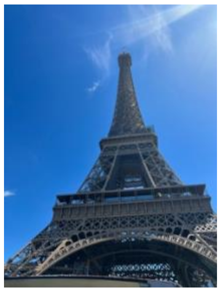
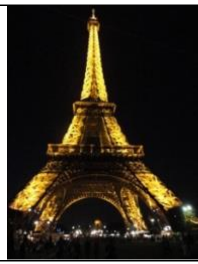
A	B	C	D	E	F	G	H
Website	**Rank**	**Visitors**	**Industry**	**Upload limit**	**Guidelines—photos must…**	**Viewing**	**Votes for usefulness**
Google	1	3.134B	Search engine	None	be related to the listing, original, not excessively stylistically altered, and family friendly.	With text review or in a separate photo gallery	Thumbs Up button
Amazon	12	522.5M	Market place	None	not include false information, be non-promotional, related to the listing, original, and avoid mature content.	With text review or in a separate photo gallery	Thumbs Up button
Booking	52	210.5M	Travel	20	be related to the listing, original, useful, and family friendly.	With text review or in a separate photo gallery	Thumps Up or Thumps Down button
Ali Express	55	165.1M	Market place	5	not contain false, inappropriate, or sensitive information and must be non-discriminatory.	With text review only	Yes or No button
Etsy	67	160.3M	Shopping	None	be original, not include false information, and avoid graphic, obscene or mature imagery.	With text review only	Thumbs Up button
Trip Advisor	281	95.07M	Travel/ hospitality	None	be related to the listing, original, high quality, and family friendly.	With text review or in a separate photo gallery	Thumbs Up button
Yelp	202	87.46M	Travel / hospitality	None	be related to the listing, original, and family friendly.	With text review or in a separate photo gallery	Useful, Funny and Cool buttons

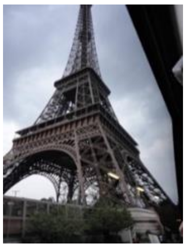
In all platforms, photos must be uploaded at the time of review, and users are asked to ensure consumer privacy. Websites Statistics and Industry Information from Pro.Similarweb.com are from September 2021. https://pro.similarweb.com/#/research/home. Date accessed: June 11, 2022.

[8] https://baymard.com/blog/allow-reviewers-to-upload-images

Our work documents a distinct effect of photos that operates above and beyond the information that they provide. We argue and find that user-generated photos are often *nonprobative*, in that they do not support the reviewer's specific claims, nor even their attitude valence. For example, the Eiffel Tower has over 100,000 user-generated photos on TripAdvisor. In Table 2, we present a sample of these photos, which show little correspondence to their attached ratings. Arguably, most tourists to Paris would find the content of these photos to be redundant and uninformative. We argue that despite these limitations, non-probative photos serve as *pseudo-evidence* (Newman, Garry, Bernstein, Kantner, and Lindsay 2012); they signal the reviewer's confidence in their review, and readers experience this as greater review value.

Table 2: Reviews of Eiffel Tower from TripAdvisor

Rating	Reviewer's text	Reviewer's photo
5 Star Reviews	The Eiffel Tower offers a lot more than you imagine. Great experience the que wasnt that long 1:30 from the moment we arrived to standing on top of the Eiffel Tower isn't too bad in my opinion. The view at the top is phenomenal	
	Eiffel Tower amazement. This was a great experience. Pay to ride the elevator to the 3rd deck and view. Champagne at top but way overpriced. Views are great. Take elevator down to second floor and enjoy cafe and such. Pictures were not worth it after excitement wore off. Walk down stairs from 2nd floor to bottom does take about 10 minutes and is enjoyable.	
	Paris Icon. Visiting the Eiffel Tower in Paris is a must do. We ate at the restaurant on the first floor with great views and food was better than expected. There is long lines though and you must have tickets beforehand.	
1 Star Reviews	Do not have high expectations. I want to tell ya' all that this place is not as good as you imagine it to be. Very bad place and views are that of craggy skyline.	

	Below expectation. I was disappointed. This place is up there is dashing your expectations big time. Not totally what you envisioned in your head before you come. In getting in, views, access, not good at all.	
	Complications from too many people. There are too many people in here. It is very likely that complications could arise from having this many people here. It is filthy place.	

NOTE. Reviews with photos were selected as the most recent three, and in which people were not the focus of the photo (dated June-July 2022).

Our research offers two main theoretical contributions. First, research on online reviews has studied linguistic characteristics of review text (Kronrod and Danziger 2013; Moore and Lefreniere 2019; Schellekens, Verlegh, and Smidts 2010) and how consumers interpret ratings distributions (He and Bond 2015; Mudambi and Schuff 2010). Relatively less work has examined the role of user-generated photos in online reviews (for exceptions, see Ma, Xiang, Du, and Fan 2018; Yang, Shin, Joun, and Koo 2017). To address this gap, we test the value of user-generated photos while ruling out potential alternative explanations for the effects.

Second, we contribute to research on the consumer psychology of photos (e.g., Diehl, Zauberman, and Barasch 2016; Henkel 2014; Taylor 2020), which has studied why consumers take photos during experiences and the impact on their own engagement. We add to this literature by studying the social transmission of photography and the impact on photo *viewers.* Our theory synthesizes research on signaling effects (Kihlstrom and Riordan 1984) and nonprobative photos (Newman et al. 2012) to show the signaling value of photos.

This work has implications for review platforms and their policies and actions on user-generated photos. Overall, we find that photos make reviews more valuable—an effect that applies to individual reviews, collections of reviews, and review sites that allow photo uploads. Importantly and distinct from past work, we find that photos also drive greater choice likelihood and decision confidence, even when they contain no diagnostic information. Review platforms can draw upon these findings to design platforms in ways that encourage more user-generated photos of a broader range.

Theoretical Background

User-Generated Photos in Online Reviews

Online reviews are highly influential in many business domains, affecting brand image and attitudes (Jalilvand and Samiei 2012; Shihab and Putri 2019) and product sales (e.g., Baker, Donthu, and Kumar 2016; Chevalier and Mayzlin 2006; Vana and Lambrecht 2021). Research on online reviews has mainly focused on ratings (e.g., De Langhe, Fernbach, and Lichtenstein 2016; He and Bond 2015; Watson, Ghosh, and Trusov 2018) and review text (e.g., Kronrod and Danziger 2013; Filieri 2016; Rocklage and Fazio 2020; Zhang, Zhao, Cheung, and Lee 2014). Relatively less is known about user-generated photos.

A survey by Bazaarvoice (2020) found that having access to user-generated photos increased purchase likelihood for 62% of consumers. Yet, there are also tradeoffs that limit practitioners' enthusiasm for collecting user-generated photos. Asking reviewers to post a photo with their review adds friction and risks of misuse and data privacy. For these and other reasons, some e-commerce sites do not collect user-generated photos (Baymard Institute 2020), and many others deemphasize photos on their review sites.

Some academic work has studied the value of user-generated photos (e.g., Zinko, Burgh-Woodman, Furner, and Kim 2021) and tested their usefulness on specific platforms (e.g., TripAdvisor; Park, Sutherland, and Lee 2021; Hlee, Lee, Yang, and Koo 2019). Yet, less is known about the generalizability of these findings and the underlying mechanisms for these effects. Past work has found that user-generated photos assist readers by illustrating and augmenting review text (Ma et al. 2018). We acknowledge that user-generated photos can offer supportive information, but we explore a distinct process related to the psychology of photos, which has different implications for practitioners.

Psychology of Photos

Recent research has explored the role of photos in consumption experiences and product usage occasions (e.g., Barasch, Diehl, Silverman, and Zauberman 2017; Barasch, Zauberman, and Diehl 2018; Henkel 2014). According to this work, taking photos during experiences enhances consumers' enjoyment, immersion, and memory retention (Diehl and Zauberman 2022; Diehl, Zauberman, and Barasch 2016; Tonietto and Barasch 2020). Thus, consumers may take photos of products, services, and experiences for various reasons.

However, research has been relatively silent on how photos taken by consumers affect others in social influence settings, such as on online review platforms. Consumers perceive photos to be easier to process than text (Alba and Hutchinson 1987; Babin and Burns 1997; Edell and Staelin 1983). In turn, photos included with product information may affect others' attitudes, beliefs, and inferences by providing visual evidence (e.g., Meyers-Levy and Peracchio 1992; Miniard, Bhatla, Lord, Dickson, and Unnava 1991; Mitchell and Olson 1981; Peracchio and Meyers- Levy 1994; Smith 1991)

Yet, visual evidence may not be the only nor dominant benefit of posting photos with reviews. Another line of research on the psychology of photos has shown that photos can boost confidence even when they offer no probative value (Cardwell, Henkel, Garry, Newman, and Foster 2016; Newman and Zhang 2020). For example, in a study on political trivia, seeing a photo of a former political leader led people to have greater confidence in the claim that he is living and in the claim that he is deceased (Newman, et al. 2012). In this example, the photo depicts the politician's appearance, but it does not reveal his mortality status. The politician's photo is *nonprobative*, in that it can neither confirm nor disconfirm either claim.

Within online reviews, we posit that user-generated photos may also be nonprobative. Reviewers may lack the skill or incentives to capture and communicate sentiments through photos (Marder, Erz, Angell, and Plangger 2021). In experiential categories like meals and vacations, consumers may have taken photos for other purposes, such as to collect memories or to post on social media (e.g., Diehl et al. 2016). Many photos that consumers post for these domains were taken incidentally during consumption and uploaded later.

Research has shown that presenting related, but nonprobative photos alongside statements promotes "truthiness"—the sense that a statement is true (Fenn, Newman, Pezdek, and Garry 2013; Newman et al. 2012; Newman, Garry, Unkelbach, Bernstein, Lindsay, Nash 2015; Newman, Garry, Bernstein, Kantner, and Lindsay 2012). This effect has been observed in domains like beliefs in general knowledge claims (Newman, et al. 2015) and predictions about the future (Newman, Azad, Lindsay, and Garry 2018). Non-probative photos can also trigger false memories about past experiences (Cardwell, et al. 2016), and alter post-experience product evaluations (Cardwell, Newman, Garry, Mantonakis, and Beckett 2017).

To explain these effects, research has supported an ease of information processing account (Newman et al. 2015; Wilson and Westerman 2018), in which photos give an increased sense of familiarity that may be misconstrued as an indication of truth (Newman, et al. 2015; Newman and Zhang 2020; Zhang, Newman, and Schwartz 2021). However, we believe that nonprobative photos may affect decision-making differently in online review contexts, for three reasons. First, in online review contexts, consumers generally have marketer-provided photos to familiarize themselves with the product, service, or experience, which softens any benefit from adding familiarity. Second, familiarity may be beneficial for verification of information, but evaluating online reviews involves far more subjectivity. Third, photos with online reviews are also socially transmitted and have persuasive content. Thus, consumers may derive information from another person's act of posting a photo, and not just the photo itself.

Signaling Effects and Confidence

Per our reasoning above, we argue that user-generated photos within online reviews may not ease readers' processing of review text. If so, why would nonprobative photos increase the perceived value of a review? We argue that photos signal the reviewer's confidence in their evaluation. This contention draws on signaling theory (Boulding and Kirmani 1993; Akerlof 1970), an account for how consumers rely on signals in their decision-making.

Past work on signal effects in advertising has shown that consumers perceive advertising expenditures as signals of a firm's confidence in its claims (Nelson 1974; Milgrom and Roberts 1986; Hertzendorf 1993). Similarly, linguistic research has studied how observers use signals such as speech rate (Street, Brady, and Lee 1984) and eye gaze (Ridgeway, Berger, and Smith 1985) to determine the confidence of speakers. In line with these past findings, we posit that user-generated photos act as signals, reflecting the confidence of the reviewer. Photos may be perceived as less manipulated and more realistic (Sundar 2008). Additionally, posting photos involves self-presentation;

photos and their attached reviews may be salient and noticed by readers (Bigne, Chatzipanagiotou, and Ruiz 2020; Li and Xie 2020).

By interpreting this signal as a cue of reviewer confidence, readers may in turn, generate pseudo-evidence; they may sense that a photo supports the review text (Cardwell et al., 2017; Derksen, Giroux, Connolly, Newman, and Bernstein 2020; Fenn, et al. 2013; Fenn, Ramsay, Kantner, Pezdek and Abed 2015; Newman and Zhang 2020; Zhang, Newman, and Schwarz 2021). As consumers often rely on such simple cues when evaluating online content (Dunaway, Searles, Sui, and Paul 2018), we argue that the mere presence of photos generates pseudo-evidence in support of the review text, which enhances the review's perceived value. Thus, we hypothesize:

H1: Including photos alongside review text and ratings increases readers' perceived value of the review, even when the photo does not support the reviewer's evaluation or text-based claims.

Additionally, we posit that user-generated photos generate confidence on behalf of the reader. Reader confidence refers to one's conviction regarding their assessment of the review's evaluation and content (Tormala 2016). As we discussed earlier, the inclusion of photos can enhance trust (Newman and Zhang 2020) and familiarity (Zhang, et al. 2020). The subsequent pseudo-evidence can boost readers' confidence. Research has demonstrated that accumulating nondiagnostic information can increase confidence (e.g., Bell and Loftus 1989; Tsai, Klayman, and Hastie 2008). Thus, we argue that photos, compared to text alone, will result in readers feeling more confident in the reviewer's evaluation.

H2: Readers perceive that reviewers who post photos alongside their text and ratings signal confidence in their review. This in turn increases reader's own confidence in the reviewer's evaluation and raises review value.

Overview of Studies

We test these hypotheses across seven studies. Study 1 uses field data from Yelp to show that reviews with photos are perceived as more valuable, and we address alternative explanations. Studies 2 and 3 support our account that even nonprobative boost review value. Studies 4a and 4b provide initial evidence for our proposed confidence signaling mechanism, and Study 5 extends this with a mediation analysis. Finally, in Study 6 we moderate the effect of photos on review value by altering the cue attached to photos.

Study 1: Value of photos in Yelp Restaurant Reviews

Study 1 was designed to test our basic effect, that reviews with photos are perceived as more valuable. While some past work has observed this relationship (e.g., Hlee, et al. 2019, Li, Law, Xie, and Wang 2021), Study 1 enhances those findings through the analysis of real-world data and addresses alternative explanations. We analyzed 15,461 Yelp reviews, sampled from 75 restaurants in three U.S. towns. Our independent variable was the inclusion of photos with a review, and our dependent variable was the number of 'useful' votes obtained. To test competing accounts, we also measured reviewer, restaurant, and review characteristics, and we analyzed the review text with the software LIWC[9] (Pennebaker, Francis, and Booth 2001). Table 3 has a description of several covariates and their correlations with votes of 'useful.'

[9] Linguistic Inquiry and Word Count (LIWC) is a popular software package for studying the emotional, cognitive, structural, and process dimensions in text analysis. LIWC produces summary measures of Analytical, Clout, Emotional Tone, and Authenticity, described in Table 3. The software counts the number of words in the sampled text from a preset LIWC lexicon (e.g., words for positive valence, words indicating analytical thinking, etc.). The index values range from 0 to 100 (i.e., weak to strong presence in the sampled text).

Table3: Key variables used in Study 1: Yelp Field Data Analysis

Type of variable:	Variable	Description	Mean (SD)	Correlation with votes of 'useful' *r*(15473) =…
Covariate: Reviewer characteristics	Reviewer's average	Average of all of the reviewer's restaurant ratings	66.5 (114.2)	0.23***
	Reviewer's count	Total # of reviews on Yelp written by the reviewer	77.82 (198.5)	0.2***
	Reviewer's friends	Total # of friends of the reviewer on the Yelp platform	55.92 (212.8)	0.24***
Covariate: Restaurant characteristics	Restaurant's average	Average of all the target restaurant's ratings	4.04 (0.41)	-0.05***
	Restaurant's count	Total # of reviews for the restaurant on Yelp	282.58 (143.95)	-0.01
Covariate: Review text, using LIWC	Analytic	% algorithm: use of formal, logical, hierarchical thinking patterns	57.67 (28.38)	0.01
	Clout	% algorithm: social status, confidence, leadership in writing	49.24 (27.33)	-0.03***
	Authentic	% algorithm: personal, humble, vulnerable writing	47.07 (33.83)	0.07***
	Tone	% algorithm: net positivity positive emotion minus negative emotion	80.66 (29.67)	-0.19***
	Word count	Word count of the review	55.32 (44.9)	0.06***
Covariate: Review rating	Star rating	Rating of the restaurant by the reviewer on a 5-star scale	4.05 (1.17)	-0.21***
Covariate: Review age	Days since post	# of days elapsed since the review was posted (at time of data collection)	1975.44 (1181.43)	-0.1***
Independent variable	Photos	# of photos posted with the review	0.21 (0.64)	0.13***
Dependent variable	Useful	Votes of 'useful' for the review by readers	0.9 (1.74)	1.0

Results

Reviews with 1+ user-generated photos were *N* = 1,930 (12.5% of the data). Specifically, 1,040 had one photo, 416 had two photos, and 474 had three photos (the maximum). Our key predictor was the presence or absence of a photo (binary measure), but results replicate when examining photo count as a continuous measure. The dependent measure, votes of 'useful' (*M* = 0.90, *SD* = 1.74), had a Pareto distribution with the 20% most useful reviews garnering 76% of the 'useful' votes. Most reviews (58%) received no useful votes

As expected, reviews with 1+ photos received significantly more useful votes (M = 1.36, SD = 2.5) than those without photos (M = 0.84, SD = 1.6; $t(15459) = 12.30$, $p < .0001$). Receiving useful votes was more likely with a photo (54%) versus without (40%; $z = 11.2$, $p < .001$). Most reviews (75.86%) were positive (4+ stars out of 5)—86.7% of reviews with a photo and 65% of reviews without a photo. The effect of photos on votes of 'useful' held separately for positive (M_{Photo}= 1.24, SD = 2.30 vs. $M_{No\text{-}photo}$= 0.62, SD = 1.19; $t(1823)$= 10.82, $p < .0001$) and negative reviews (M_{Photo}= 3.22, SD = 4.94 vs. $M_{No\text{-}photo}$=.89, SD = 2.97; $t(8722)$= 2.47, $p < .015$).

We subjected the count of useful votes to a Poisson regression, where our key predictor, presence of a photo, was significant (b=0.35, p<.0001). To test the robustness of this effect, we conducted further analyses in which we controlled for characteristics of the review, reviewer, and target restaurant, see Table 4. The effect of photos on votes of useful remained significant (p<.0001) when controlling for each factor.

Discussion

Study 1 offers empirical support, through real-world data, for the perceived value of user-generated photos. Indeed, results suggest that using photos, rather than text-alone, is worthwhile to boost the value of review content. While such a result aligns with some research in practitioner (Kats 2021; Sims 2020) and academic journals (e.g., Hlee, et al. 2019, Li, et al. 2021), we show that our effect holds even after accounting for confounding factors, including verifiability, trust, authenticity, and valence of reviews and reviewers.

Table 4: Alternative Explanations

Variable Source	Alternative explanation	Covariates in the analysis
Review	Reviews attached to photos are more positive/ negative	Star rating, LIWC ‘Tone’
	Reviews attached to photos are expressed more emotionally	LIWC: ‘Tone,’ 'positive emotion,' 'negative emotion'
	Reviews attached to photos contain more information	Word count of the review
	Reviews attached to photos are more verifiable/ authentic	LIWC: 'Authentic'
	Reviews attached to photos are better expressed / more cogent	LIWC: ‘Analytic,’ ‘insight’, ‘cognitive processing’
	Reviews attached to photos had been posted earlier or more recently	‘Review age’: # of days elapsed since the review
Reviewer	Reviewers who post photos are more experienced	# of reviews on Yelp by the reviewer, # of days since joining Yelp
	Reviews who post photos are more positive/ negative	Average of past star ratings by reviewer
	Reviewers who post photos are more popular, trustworthy	# of 'friends' on the Yelp platform
	Reviewers who post photos are more local	U.S. state of the reviewer (vs. state of the target restaurant)
Restaurant	Restaurants with more photos are more positive/ negative	Average star rating of the restaurant
	Restaurants with more photos are more popular	Total # of reviews of the restaurant

Study 2: Matching photos with Trip Advisor Review Text

In Study 2, we aim to show that photos with online reviews may be nonprobative, in that they may not support the review’s claims or even the review’s attitude valence. To that end, we hypothesize that when photos and their attached review text are separated, people would be unable to consistently match them. Using actual TripAdvisor reviews, we asked participants in our study to try to identify the review text attached to a photo, where a decoy review's text differed in content and valence from the correct text.

Method

Participants from mTurk (N = 110, M_{age} = 42, SD = 13.66; 66.2% female) completed a pre-registered experiment (AsPredicted#84339). They were asked to match TripAdvisor photos to their review text over 10 trials, see Methodological Details Appendix F. In all trials, the photo was sourced from a positive review (4 or 5 stars). In five trials, participants had to identify which of two real reviews included this photo, where both options were positive. In the other five trials, the correct text was positive, but the decoy text was negative, sourced from a review of 1 or 2 stars. For our stimuli, we selected the most recent consumer-posted photos and reviews for five popular London landmarks. For each review pair, participants were asked to "indicate which of the reviews you think was uploaded with the photo." We combined the responses from the 10 trials to form an accuracy level between 0% and 100%. Following this, four simple photo identification tasks were employed as an attention check, which four participants failed, resulting in 106 usable responses.

Results

For each participant, we computed a score of accuracy (0%-100%). A hypothesis test for accuracy found that participants were not able to successfully match the photo to the correct text better than 50% of the time (M = 52%, SD = 15.13%; t(110) = 1.4, p = .17). See Table 5 for results by trial.

Table 5: Study 2 Results

Trial	London landmark	Valence of decoy review's text	Accuracy rate
1	The Shard	Positive	86.24%
2	The Shard	Negative	14.68%
3	London Eye	Positive	62.39%
4	London Eye	Negative	55.05%
5	Buckingham palace	Positive	66.06%
6	Buckingham palace	Negative	47.71%
7	Tower Bridge	Positive	61.47%
8	Tower Bridge	Negative	33.03%
9	Oxford Street	Positive	54.13%
10	Oxford Street	Negative	45.87%

NOTE. In each trial, participants indicated which of two reviews was attached to a photo of a London landmark. Trials appeared in random order and text decoys were counter balanced left-right.

Discussion

Study 2 demonstrates that many real-world photos with reviews do not illustrate the attached, text-based claims. The decoy reviews had different content and valence from the source review. Still, participants were unable to match the photo to the correct text better than chance guessing. Our argument is not that all photos are uninformative; for example, user-generated photos may depict features and defects or simply verify that the user has purchased the product. Rather, our argument is that even when user-generated photos do not achieve these aims, they still boost review value for reasons that we explore in our subsequent studies.

Study 3: Pizza Restaurant Reviews

In Study 3, we conduct a controlled scenario-based experiment to test our focal hypothesis that the inclusion of a photo enhances review value, even if the photo is nonprobative. We test this account by using the same photo with different review texts, and we predict that photos will have a similar effect on review value in either case.

Method

Participants from Prolific (N = 200, M_{age} = 40.38, SD = 10.66; 50% female) completed a pre-registered experiment (AsPredicted#102740) that had a 2 (photo: probative vs. nonprobative) x 2 (review valence: positive vs. negative) between-subjects design.

Participants were asked to consult two reviews of a pizza restaurant. The two reviewers' text and ratings were positive (4 stars out of 5) for half the participants and negative (2 stars out of 5) for the other half of participants. In each valence condition, one of the reviewers posted a photo with their review. In the probative cell, the photo matched the review text; for example, the negative review referred to messy pizza that fell apart, and the accompanying photo showed a messy pizza box. In the nonprobative cell, the photo was attached to a different text; for example, the messy pizza box photo was attached to a review that referred to slow staff. These reviews and photos were pre-tested to ensure that they differed in how probative the photo was for the review text. See Appendix G for stimuli and pre-test results. After seeing the pair of reviews, participants responded to our dependent measure. They rated the relative helpfulness of these posts on a 2-item (α = .92), 7-point scale 'helpful' and 'useful' (1 = "Much more Post A" to 7 = "Much more Post B").

Results

The review helpfulness measure was subjected to an ANOVA with independent factors of photo and valence conditions. As expected, the main effects and the interaction were not significant (all $F(1, 198) < 2.34$, $p > .13$). Planned contrasts revealed that the reviews with a photo were perceived as equally helpful, whether the photo was probative or nonprobative. This occurred both in positive reviews ($M_{nonprobative}$ =5.69, SD = 1.32 vs. $M_{probative}$=5.61, SD = 1.59; $t(196) = -0.266$, $p = .79$) and in negative reviews ($M_{nonprobative}$= 5.11, SD = 1.71 vs. $M_{probative}$ = 5.67, SD = 1.35; $t(196) = -1.877$, $p = .07$). In the latter, the difference between conditions approached significance, but the effect is in the opposite direction of a rival prediction, that people would see probative photos as more helpful.

Importantly, the means in all four cells were significantly different from the scale mid-point ($t(49) > 4.61$; $p < .0001$), suggesting that the review with a photo was perceived as more helpful than the review without the photo, even if the photo was unrelated to the reviewer's claims.

Discussion

Study 3 offers additional support for the role of photos on review platforms. We show that photos acted to enhance review helpfulness regardless of review valence. Moreover, our results show that photos remain helpful despite being non-probative, as the same photo was able to enhance review helpfulness in both review valence conditions. This finding further demonstrates that user-generated photos can improve review value, even if they offer little to no supportive evidence.

Study 4: Confidence in BBQ Reviews

In Study 4, we test our theory that nonprobative photos improve review value by boosting confidence. We conducted two studies on barbecue ("BBQ") pits: Study 4a on reviews that were more positive and Study 4b on reviews that were less positive. Across both studies, we use the same photos to support the reviews. In this study, participants respond to a gallery of photos and a collection of reviews, rather than to an individual review, to show how the phenomenon impacts broader outcomes.

Method

In both studies, participants were asked to evaluate a pair of BBQ pits, options A and B, that were similar in price, functionality, materials, and features. See Appendix H for stimuli. All participants also saw marketer-provided photos of both BBQ pits, which were from real-world products listed on Amazon. Thus, all participants knew what the BBQ pit looked like generally, having seen professional photos of the product.

We pre-tested the reviews on mTurk (N = 58), see Appendix H. Each review was ~10 words long, adapted from real reviews on Amazon. Pre-test participants rated 12 reviews for how helpful they were versus the typical online review on a 5-point scale (1 = "much less helpful", 5 = "much more helpful"). All reviews were perceived to be moderately helpful (M = 3.52, SD = 1.01; range: 2.98 to 3.87). They also rated each review for its valence from -5 (negative) to 0 (neutral) to +5 (positive). We identified six reviews that were positive (M = 3.01, SD = 1.90) and six reviews that could be classified as negative or neutral (M = -0.85, SD = 2.88).

In the main study, participants were shown five counterbalanced consumer reviews for each BBQ pit option. For half of the participants, both options did not have any photos attached to their reviews. For the other half of participants, BBQ B had a gallery of three user-generated photos, which were sourced from actual Amazon reviews of that BBQ pit. In this condition, BBQ A had no photos from reviewers. After viewing the set of reviews, participants responded to the measures in which they compared the two BBQ pit options, order counterbalanced. Further details are provided separately for each study below.

Study 4a: More Positive BBQ Reviews

Participants from mTurk (N = 127, M_{age}= 41, SD = 13.05; 59% female) completed a pre-registered study (AsPredicted#69639) that had a 2-cell (photo condition: photo vs. no photo) between-subjects design. In Study 4a, for each option, three reviews were positive, and two reviews were negative or neutral.

After reading the reviews, participants responded to two measures. First, they rated their relative likelihood of selecting BBQ A versus B on a 7-point scale (1="Much more likely A", 7 = "Much more likely B"). They then rated how confident they felt about the two options on a 3-item (α = .91), 7-point scale (1 = "A lot more confident in A", 7 = "A lot more confident in B"). They rated their confidence in which option was (1) more popular, (2) recommended by more customers, and (3) had more satisfied customers.

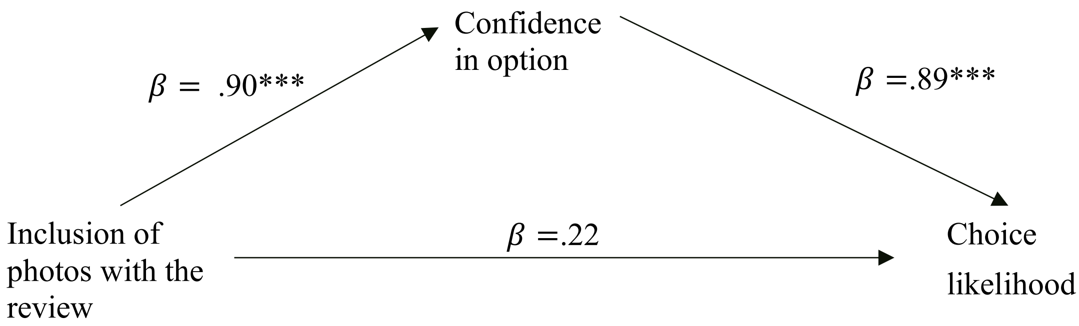
Study 4b: Less Positive BBQ Reviews

Participants from mTurk (N = 126, M_{age} = 37.1, SD = 12.19; 52% female) completed a pre-registered study (AsPredicted #70931) that had a 2-cell (photo condition: photo vs. no photo) between-subjects design. The study design was very similar to Study 4a, with two differences. First, for both BBQ pits, three reviews were negative or neutral, and only two were positive. The three photos of BBQ B used in this study were identical to those presented in Study 4a. Second, we adapted the relative choice likelihood measure to refer to rejecting the options, because both BBQ pits had more negative than positive reviews, which would make both options undesirable for purchase (see Appendix H). However, for our analysis, we focus on the negative confidence measure: On a 3-item (α = .91), 7-point scale (1 = "A lot less confident in A", 7 = "A lot less confident in B"), participants rated which of the two undesirable options they felt *less* confident about in terms of its popularity, recommendation by customers, and customer satisfaction.

Results

In Study 4a (more positive reviews), the relative choice of BBQ B was greater when it included photos (M = 4.73, SD = 1.57) versus when it did not have photos (M = 3.70, SD = 1.60; $t(125)$= 3.64, $p <$ .001). Relative confidence in BBQ B's popularity was also greater with photos (M_{Photo}= 4.70, SD = 1.48 vs. $M_{No\text{-}photo}$= 3.80, SD = 1.24; $t(125)$ = 3.73, $p <$ .001). We tested the mediating role of confidence in this effect: A bootstrap-based analysis (5,000 resamples) indicated a significant indirect effect and complete mediation (β= 1.027, SE=.28.; 95% CI=[.479 to 1.57.]), see Figure. As hypothesized, including photos with the review increased confidence in the target option, which in turn increased preference for this option.

Figure 1: Mediation model in study 4A



In Study 4b, relative confidence in BBQ B's lack of popularity was greater when it included photos (M = 4.11, SD = 1.39) versus when it did not have photos (M = 3.58, SD = 1.42; $t(124) = 2.12$, p = .04).

Discussion

Study 4A and 4B demonstrate that a gallery of user-generated photos increases consumers' sense of confidence from a set of consumer reviews. All participants had already seen professional photos of the product before encountering the reviews. Moreover, the same photo gallery enhanced consumer confidence in different directions. In more positive reviews, photos boosted confidence in product utility, whereas in negative reviews the same photos boosted confidence in product disutility. This pattern suggests that it is less likely that the user-generated photos impacted confidence by providing supportive evidence. In Study 4A, we further found that an increase in confidence mediated the effect of nonprobative photos on choices.

Study 5: Serial Mediation Confidence in Massage Gun Reviews

In Study 5, we test our proposed mechanism that photos improve review value through an increase in confidence. More specifically, we hypothesized (H2) that readers of reviews would perceive a greater signal of confidence by reviewers who post a photo with their review, and this in turn increase readers' own confidence in the reviewer's evaluation. We test this process through mediation. Study 5 also generalizes the finding to a different product category, massage guns.

Method

Participants from Prolific (*N*=160, M_{age}=39, *SD*=9.77; 50% female) completed a pre-registered study (AsPredicted #105160) that had a 2-cell (photo condition: photo vs. no photo) between-subjects design. The task instructions were similar to Study 4 but adapted to a different product category and with a single review of each option (vs. multiple reviews in Study 4). More specifically, participants were asked to consider two massage gun options TYIAUS and WESKEAN, that were similar in price, functionality, materials, and features. See Appendix I for stimuli. Before seeing the review, all participants were shown marketer-provided photos of both massage gun options from real-world products listed on Amazon to ensure a degree of product familiarity.

Next, participants were shown a single, positive consumer review for each massage gun option. Half of the participants saw a review without user-generated photos for both product options. The other half of participants saw a review for TYIAUS with a gallery of three user-generated photos that were sourced from real-world Amazon consumer reviews. Moreover, in this condition, the WESKEAN massage-gun option had no user-generated photos. Each massage gun option post was assigned a reviewer author name "Adam" and "Bob", which were counterbalanced. After viewing the pair of reviews, participants responded to the measures in which they compared the two massage gun options with a order counterbalanced.

After reading the reviews, participants responded to three measures. First, they rated their relative likelihood of selecting massage gun WESKEAN versus TYIAUS on a 7-point scale (1="Much more likely WESKEAN", 7="Much more likely TYIAUS"). They then assessed each reviewer's confidence in their post on the massage gun on a 3-item (α=.90), 100-point, unmarked slider scale[10]. They rated which reviewer (1) signaled more confidence, (2) displayed more belief in their post, and (3) showed that they stood by their post more. We converted the scale to an index indicating degree

[10] We used a 100-point scale in this first measure to address potential common method bias.

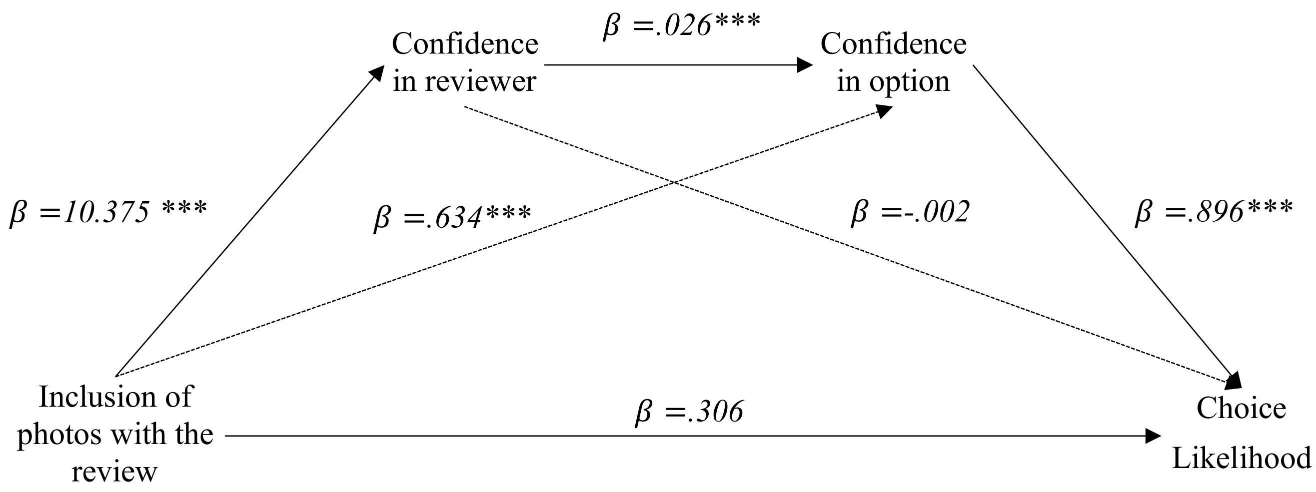
of signaled confidence by the reviewer who posted for TYIAUS, Bob. Finally, participants rated how confident they felt about the two options on a 3-item (α=.95), 7-point scale (1= "A lot more confident in WESKEAN", 7="A lot more confident in TYIAUS"). They rated their confidence in which option was (1) more popular, (2) recommended by more customers, and (3) had more satisfied customers.

Results

The relative choice of massage gun TYIAUS was greater when it included photos (M=5.10, SD=1.51) versus when it did not have photos (M=4.00, SD=1.57; t(158)=4.52, p<.001). Relative perceived confidence of reviewer Bob was also greater with photos (M_{Photo}=58.4, SD=15.87 vs. $M_{No\text{-}photo}$=48.1, SD=16.02 ; t(158)= 4.11, p<.001). Further, self-rated relative confidence in TYIAUS' popularity was also greater with photos (M_{Photo}=4.67, SD=1.10 vs. $M_{No\text{-}photo}$=3.76, SD=1.04; t(158)= 5.37, p<.001).

We tested our proposed process through a serial mediation model. According to this model, inclusion of a photo with the review (X) increases readers' sense that the reviewer has signaled confidence in their review (M1), which increases self-rated confidence in the product (M2), which ultimately increases choice likelihood for the product (Y). A bootstrap-based analysis (5,000 resamples) indicated a significant indirect effect (β=0.244, SE=.01, Z=2.47; 95% CI=[.094 to 0.495.], p=.014). This pathway fully accounted for the overall impact of *Photos* on *relative choice* with the direct effect being insignificant (β=0.306, SE=.264, Z=1.16; 95% CI=[-.166 to 0.856.], p>.10), see Figure 2.

Figure 2: Serial Mediation Model in Study 5



Discussion

Study 5 demonstrates that a single review accompanied by a gallery of user-generated photos increases consumers' sense that the reviewer is confident in their review and consequently increases consumers' own confidence in their evaluation of an option. Thus, this study supports our proposed process through serial mediation. Like in Study 4, all participants saw professional photos of the product before encountering the reviews, but the inclusion of user-generated photos increased choice likelihood, nonetheless. Study 5 further supports the idea that, in addition to claims of product quality, consumers integrate cues of confidence, related to the reliability of the review, into their decision making.

General Discussion

Online reviews affect consumers' evaluations and spending (Baker, Donthu, and Kumar 2016; Chevalier and Mayzlin 2006; Jalilvand and Samiei, 2012; Shihab and Putri 2019). However, user-generated photos in online reviews have received scant attention in academic literature (for exceptions see Park, et al. 2021; Hlee et al., 2019; Zinko, et al. 2021). Our work attempts to fill this gap and identifies a novel pathway by which photos affect review value.

Through real-world Yelp data, we show that reviews with photos, compared to text alone, are more valuable. Our study also show that user-generated photos also boost choice likelihood for positively rated options. Further, we show that the value of photos stems from their use as a signal. Drawing on signaling theory (Nelson 1974; Milgrom and Roberts 1986; Hertzendorf 1993) and literature on nonprobative photos (Newman, et al. 2012; Newman and Zhang, 2020), we find that photos signal reviewer confidence, which can lead to enhanced reader confidence and subsequently review value.

Theoretical Implications

While online review research has focused on text and rating characteristics (Kronrod and Danziger 2013; Moore and Lefreniere 2019; Schellekens, Verlegh, and Smidts, 2010), our research contributes by demonstrating the value brought by user-generated photos on review platforms while considering alternative explanations. We also find that their value holds, despite a lack of informational value. Therefore, our results shed light on user-generated photos within this context, by demonstrating their ability to boost review value through mere inclusion, adding to prior literature that has documented alternative means by which review value can be enhanced, including reviewer profile dynamics (Karimi and Wang, 2017), text characteristics (Kronrod and Danziger 2013; Filieri 2016; Rocklage and Fazio 2020; Zhang, et al. 2014) or website design (Aerts, Smits and Verlegh 2017).

This research also contributes to past signaling literature (Boulding and Kirmani 1993; Akerlof 1970; Spence 1974) in two ways. First, extending advertising research (Boulding and Kirmani 1993; Erdem, Keane, and Sun 2008; Kihlstrom and Riordan 1984) that has documented the use of expenditure and frequency to signal quality, we find that photos act to provide a visual cue that that the reviewer is confident in their evaluation. Second, we argue that this signal generates pseudo-evidence through a belief that the photo must offer supportive information, leading to an increase in reader confidence and thus, review value. This finding offers insights into the social influence of photos on viewers, in that visual information has an influence over, not only review value or purchase intent, but also on one's confidence. Moreover, our findings suggest that consumers may not seek information from reviews alone (e.g., Fillieri 2015). In that regard, we extend prior literature that has explored review value from an information acquisition perspective (e.g., Chen and Xie 2008; Liu and Park 2015; Mundambi and Schuff 2010), by showing that review value may be influenced based on the extent to which one feels confident after viewing.

Finally, our work is consistent with prior literature on non-probative photos (Fenn, et al. 2019; Newman, et al. 2012; Newman and Zhang 2020), in that a user-generated photo, despite its informational value, can have a significant downstream impact. While scholars have employed photos to enhance claim belief (e.g., Zhang, Newman, and Schwarz 2021) or willingness to share information (Fenn, et al. 2019), we find that, a user-generated photo influences both reader confidence, as well as review value and consumer choice. Additionally, our results also provide insights into the construct. Although we do not discount the fluency perspective, our methods and results point towards an illusion of evidence through the photo heuristic and the generation of pseudo-evidence. First, the photos presented were conversationally and conceptually related to the text, thus allowing respondents to view the photo as relevant and thus, useful. Such perceptions of relevancy can result in the creation of pseudo-evidence. Second, as our work aims to test confidence, rather than ease of processing, our simulated online review platform did not differ in fluency or imaginability from review to review.

Managerial Implications

The findings of our research offer several implications for practitioners. First, while there exist numerous photo guidelines that should be adhered to when uploading, it may be better to deemphasize the requirement on photo informativeness and instead invite photos more inclusively.

Additionally, our findings suggest that consumers may use online reviews for more than just information. When confidence is particularly relevant for consumer decision-making, marketers may wish to include user-generated photos in their communications, allowing consumers to reach a cognitive threshold at a faster rate, judging that they have sufficient information to move on to the next stage of the decision-making process. Thus, the benefits are twofold. For brands, photos may facilitate quicker decision making, while review platforms may be deemed more useful to consumers.

Finally, our findings regarding the influential role of photos are supported by past literature that has documented their impact on trust (Newman, et al. 2012), recall (Cardwell, et al. 2016) and even product evaluations (Mantonakis, Cardwell, Beckett, Newman, and Garry, 2014). This influence may be beneficial to brands, but also may pose a risk to consumers. Indeed, scholars have shown that photos can lead people into believing false claims (Cardwell, et al. 2016) or misinformation (Fenn, et al. 2019). While we encourage the use of user-generated photos, we caution that their use may result in readers feeling a sense of false confidence or in one believing a claim that would otherwise be deemed untrue.

Recommendations for Future Research

We believe that our findings could be extended in future work to better understand the underlying message conveyed when consumers include photos in their communications. For example, photos with consumer complaints may signal greater annoyance (vs. text alone), and more generally, sending a photo may express greater emotionality, even if the photo contains little emotion. Further, sharing a photo may signal disclosure or intimacy. Future research may examine other messages that may be extracted from the mere act of sharing a photo.

Our Yelp data analysis reveals a large effect of photos on review value. Of course, the impact of photos on review value would depend on photo characteristics as well as the quality of other information in the review set. Our modest claim is simply that nonprobative photos may have a much larger effect on review value than one might expect. Yet, future work may compare the effect of user-generated photos versus other beneficial review features. For example, research has shown that consumers value reviews that are more recent (Lu, Wu, and Tseng 2018) and that contain diagnostic information (Filieri 2015).

Scholars may also wish to explore further moderators that may alter the relationship between photos and usefulness, such as photo quality and content (e.g., human presence). Moreover, some review platforms incorporate videos (e.g., Amazon Video Shorts). Videos may offer richer, multi-sensory depictions of products, services, and experiences, eliciting greater arousal and evaluations extremity (Togawa and Sugitani 2021). Future research may wish to explore the role of videos on online review platforms, as well as factors that may moderate their impact.

Chapter 4: General Discussion

Introduction

Studying online social influence in the context of online reviews is crucial for understanding the changing nature of consumer behaviour in the digital age. Online reviews offer methodological advantages for studying social influence and information processing, providing insights into how these phenomena operate in the digital realm. This enables researchers to develop effective marketing strategies that meet the evolving needs of consumers.

This thesis makes several theoretical contributions to the literature on online social influence and online opinion sharing. Through the examination of the impact of N on contributor judgments, this research highlights the importance of post-purchase written evaluations in shaping the opinions of potential readers. The findings of this thesis emphasize the need for a deeper understanding of the dynamics of online opinion expression and the role of their content. This work also provides preliminary evidence of how reduced sense of responsibility leads to affect-rich opinions. The results also show the limitations of using N as a signal of information and the importance of standardizing judgment criteria in online opinion expression.

This research also contributes to the field of online review research by demonstrating the impact of user-generated photos on consumer evaluations and decision making. The findings show that even photos that lack diagnostic value can enhance review value by providing a visual cue of confidence in the reviewer's evaluation. This visual cue leads to an increase in reader confidence and review value, suggesting that consumers may not seek information from reviews alone. These results also extend prior literature on non-probative photos by showing that user-generated photos can influence both reader confidence and consumer choice.

I have delved into two important aspects that contribute to our understanding of online consumer behaviour, consumer decision making, and the interplay between consumer use of technology and online information sharing. Through these two works, I address key research questions on the topics of online social influence, consumer decision making, consumer use of technology, and online reviews.

In terms of consumer decision making, chapters 2 and 3 provide insights into how consumers use online reviews to inform their choices. The findings of the first study suggest that consumers rely on online reviews in their decision-making process, while the second study highlights the role of user-generated photos in increasing confidence and review value. These findings have important implications for businesses and marketers, who can use online opinion and user-generated photos to influence consumer behaviour.

Finally, in terms of consumer use of technology and online reviews, both studies demonstrate the increasing importance of online reviews and user-generated photos in shaping consumer behaviour. They highlight the power of online reviews and photos as tools for consumers to gather information and make decisions, and they provide important insights into the role of technology in shaping consumer behaviour.

This section provides a summary and reflection of my research, makes recommendations for future work, and highlights the new findings contributed to the field of online social influence and information processing in the context of online reviews. The implications of my findings for marketing practitioners seeking to influence consumer behaviour are discussed, and future research avenues are identified.

Recommendations for Future Research

In this next section I discuss the recommendation for future research. Notably I address the question: How do online environments challenge current social influence theories, and what does the future hold for online review platforms? I also discuss the future of consumer information gathering more broadly. Given the growing presence of consumer identities online, marketers would benefit from greater understanding how interactions online can influence consumer behaviour beyond what we already know.

The existing literature on human interaction with technology often directly applies offline social influence findings to online domains, which warrants a thorough review. While most aspects of social influence remain consistent across both contexts, it is crucial to consider the unique characteristics of online environments that may alter the effects of social influence. For example, the characteristics surrounding the online context may change the fundamental assumptions underlying these effects. The nature of online communication and the potential for selective exposure to information can affect how individuals perceive and respond to social influence.

For example, it has been the norm that social influence primarily occurs in physical environments where individuals can directly observe others' attitudes, leading to group polarization (Brauer et al., 1995), group think (Janis, 1983), shared attention (Shteynberg et al., 2014), optimal distinctiveness theory (Brewer, 1991), bandwagon (Leibenstein, 1950), or diffusion of responsibility (Latané, 1981). The works covered in my thesis suggest otherwise. I demonstrate that online social influence can occur even in the absence of direct observation of others' physical attitudes and at much greater numbers, as individuals may still be influenced by virtual cues, such as text-based opinions or photos. This finding underscores the importance of examining the unique features of online social influence and addressing the gaps in current understanding.

Moreover, I contend that information available to consumers online may exert influence beyond its inherent content. The presence of individuals, manifested through comments, likes, or other online interactions, can serve both as cues and as sources of information. In this context, engagement alone can act as an additional dimension of social influence, shaping attitudes and beliefs.

A further critical aspect of online information sharing is the effort consumers exert in disseminating information. This effort, in itself, can be influential as it may signal the importance or credibility of the shared content. This perspective emphasizes the need to examine not only the content of online information but also the context and methods of sharing, as these factors jointly contribute to the overall impact of online social influence.

In light of the existing literature on human interaction with technology and its application to online social influence, several avenues for future research in consumer behaviour emerge. These include: reassessing the applicability of traditional social influence theories in online contexts; exploring the role of virtual cues in online social influence; examining the impact of online presence and engagement on social influence; and analysing the influence of consumer effort in information sharing.

Additionally, understanding the interplay between content, context, and methods of sharing is essential for a comprehensive understanding of online social influence. Addressing these research avenues will advance our knowledge of consumer behaviour in online settings and inform the development of effective strategies for navigating and leveraging social interactions in an increasingly digital world.

Where to next? Further research is needed to explore the underlying messages conveyed through user-generated content. While previous studies have examined constructs such as writer credibility and similarity, (Cheung et al., 2009; Yang et al., 2021) there is still much to be learned about the relationship between trust in the review source and the volume of opinions.

Social media platforms may consider signalling other psychological constructs that impact trust in the review source, such as expertise, and authority. Additionally, it is important to understand how these constructs interact with other variables, such as the type of product or service being reviewed, the demographics of the reviewer and the audience, and the platform on which the review is posted.

In addition to traditional text-based reviews and photos, further research is needed to investigate the potential of other mediums for gathering consumer opinions. Voice notes, for example, have become a popular means of communication on messaging apps such as WhatsApp, and they could be a valuable tool for gathering consumer opinions. Including voice notes as a tool for collecting opinions would require consideration of new literature that examines how people communicate through voice, the characteristics of voice messages that impact perception (e.g. Eisingerich et al., 2014; Flavián et al., 2023; Zarouali et al., 2021), and how to analyse and interpret voice data.

Another medium worth exploring is video. Platforms such as TikTok and Snapchat have emerged as popular channels for sharing short videos, and they could transform into new mediums for consumers to share their opinions. Visual communication means have the potential to significantly change the consumer influence process and require further exploration of the shared attention literature.

While photos are static and capture a single moment in time, video provides a dynamic and continuous visual experience that can capture the nuances of a product or service in a way that photos cannot. Video can convey more information about a product or service, such as how it works, how it looks from different angles, and how it performs in different contexts. Additionally, video allows for the inclusion of other sensory information, such as sound and motion, which can enhance the viewer's experience and provide a more immersive representation of the product or service.

It is important to understand the impact of these new mediums on consumer behaviour and how marketers can effectively leverage them to build trust and influence purchase decisions. Future research could explore the role of voice notes and video in shaping consumer attitudes, the factors that influence the effectiveness of these mediums, and how to integrate them into existing marketing strategies. By doing so, marketers can stay ahead of the curve and effectively engage with consumers using the latest communication methods.

Advancements in technology have transformed the way consumers interact with products and services and how they share their opinions. Two emerging technologies that are gaining increasing attention in the consumer behaviour literature are augmented reality (AR) and virtual reality (VR). Meta, for example has invested $36 billion into its Reality Labs division since 2019[11]. AR technology allows users to overlay virtual images onto the real world, while VR technology creates a fully immersive virtual environment that users can interact with. Both AR and VR have the potential to revolutionize the way consumers contribute their opinions and how those opinions are consumed, providing a more engaging and immersive experience for consumers. As such, they may also lead to the development of new behavioural biases among consumers. Similarly, consumers using VR to simulate real-world scenarios may develop biases based on their virtual experiences, which may not accurately reflect their actual experiences in real life.

The use AR and VR has the potential to transform the way consumers interact with products and services, which may have implications for preference construction and information processing (e.g. Alcañiz et al., 2019; Flavián et al., 2019; Zarantonello & Schmitt, 2023). While the literature has documented the impact of traditional forms of information on preference construction, such as brand image and advertising, the impact of these new technologies is not yet fully understood.

For example, contributors using AR to create videos that overlay virtual images on real-world products may process information differently than consumers viewing traditional product images.

[11] Business Insider: https://www.businessinsider.com/charts-meta-metaverse-spending-losses-reality-labs-vr-mark-zuckerberg-2022-10?r=US&IR=T

This could lead to the development of different preferences based on the information available through AR, which may not be accurately reflected in traditional preference measures.

In addition, the use of these new technologies may affect consumer expectations of product performance and perceived product quality. For example, consumers using VR to simulate real-world scenarios may develop higher expectations of product performance based on their positive virtual experiences. This could lead to a bias towards the product based on their high expectations, which may not be reflective of the actual product performance.

Another technology gaining attention is artificial intelligence, AI. The growing influence of AI on consumer behaviour and online interactions has significant implications for text reviews, particularly concerning speed, volume, and authenticity. Powerful AI models, such as OpenAI's ChatGPT for text generation and DALL-E for image synthesis, have demonstrated the potential of AI in generating coherent and contextually relevant content. While AI-generated reviews offer potential benefits, they also raise concerns about the increased prevalence of fake reviews and the challenges they pose for consumers and businesses.

Using AI for generating text reviews can enable consumers to share their opinions swiftly and efficiently. Advanced natural language processing algorithms may assist in creating well-structured, coherent, and relevant reviews, potentially increasing the volume and quality of user-generated content. However, the ease with which AI can generate reviews also presents opportunities for bad actors to create fake reviews that appear authentic and are tailored to specific platforms and existing content.

The proliferation of fake reviews can undermine trust in online review platforms and erode the value of user-generated content. As a result, consumers may become sceptical of review authenticity, leading to decreased reliance on them for decision-making. This could adversely impact feedback quality.

Therefore, it is important for researchers and marketers to consider the impact of these new technologies on preference construction and information processing, as well as consumer expectations of product performance and perceived product quality. By understanding how these technologies affect consumer behaviour, marketers can develop more effective marketing strategies that leverage these technologies to build trust and influence purchase decisions.

Conclusion

In conclusion, this thesis significantly advances our understanding of online reviews and their influence on consumer behaviour. Drawing inspiration from network social influence, as demonstrated by Christakis and Fowler (2008), this study delves into the digital transformation of consumer judgment and decision-making. It highlights the critical role that social influence plays in shaping behaviours in online environments, where readily available opinions serve as vital informational signals for consumers.

By emphasizing the dynamics of online opinion expression and the importance of visual elements, such as photos, this study offers valuable insights for marketers navigating the complex digital landscape. The thesis critically examines online social influence literature, exploring the unique characteristics of online environments that build upon and expand traditional social influence theories.

The findings of this study provide valuable insights for marketers seeking to understand consumer behaviour online and the potential of leveraging online reviews for improvements across multiple product categories. By examining the wisdom of crowds in online contexts, this thesis contributes to a deeper understanding of the aggregate opinion and its role in influencing consumer decision-making.

Moreover, this thesis underscores the importance of addressing unanswered questions and investigating the impact of emerging technologies, such as AR, VR, and AI, on consumer behaviour. It encourages continuous exploration and refinement of our understanding of online social influence and consumer behaviour in an increasingly digital world.

It is crucial to bear in mind the significance of exercising caution and critical thinking when interpreting the vast array of written and visual information found in online reviews. As we navigate our increasingly digital world, let us remain committed to refining our understanding of online behaviour and developing effective strategies for leveraging social interactions online. Until such time, *caveat lector et caveat spectator.*

Bibliography

Bibliography: Chapter 1

An, Q., Ma, Y., Du, Q., Xiang, Z., & Fan, W. (2020). Role of user-generated photos in online hotel reviews: An analytical approach. *Journal of Hospitality and Tourism Management*, *45*, 633–640.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*(2), 99–126.

Cardwell, B. A., Henkel, L. A., Garry, M., Newman, E. J., & Foster, J. L. (2016). Nonprobative photos rapidly lead people to believe claims about their own (and other people's) pasts. *Memory & Cognition*, *44*(6), 883–896. https://doi.org/10.3758/s13421-016-0603-1

Ceylan, G., Diehl, K., & Proserpio, D. (2023). EXPRESS: Words Meet Photos: When and Why Visual Content Increases Review Helpfulness. *Journal of Marketing Research*, 00222437231169711.

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, *358*(21), 2249–2258.

Fenn, E., Ramsay, N., Kantner, J., Pezdek, K., & Abed, E. (2019). Nonprobative photos increase truth, like, and share judgments in a simulated social media environment. *Journal of Applied Research in Memory and Cognition*, *8*(2), 131–138.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Frey, V., & van de Rijt, A. (2021). Social Influence Undermines the Wisdom of the Crowd in Sequential Decision Making. *Management Science*, *67*(7), 4273–4286. https://doi.org/10.1287/mnsc.2020.3713

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. https://doi.org/10.1016/j.dss.2017.06.007

Huang, L., Tan, C.-H., Ke, W., & Wei, K.-K. (2013). Comprehension and Assessment of Product Reviews: A Review-Product Congruity Proposition. *Journal of Management Information Systems*, *30*(3), 311–343. https://doi.org/10.2753/MIS0742-1222300311

Katona, Z., Zubcsek, P. P., & Sarvary, M. (2011). Network effects and personal influences: The diffusion of an online social network. *Journal of Marketing Research*, *48*(3), 425–443.

Mantonakis, A., Cardwell, B., Beckett, R., Newman, E., & Garry, M. (2014). *The mere presence of a photo on a product label can change taste perception (Working paper)*.

Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review*, *19*(5), 969–974. https://doi.org/10.3758/s13423-012-0292-0

Park, D.-H., & Kim, S. (2008). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, *7*(4), 399–410. https://doi.org/10.1016/j.elerap.2007.12.001

Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, *67*, 78–89.

Zinko, R., de Burgh-Woodman, H., Furner, Z. Z., & Kim, S. J. (2021). Seeing is believing: The effects of images on trust and purchase intent in eWoM for hedonic and utilitarian products. *Journal of Organizational and End User Computing (JOEUC)*, *33*(2), 85–104.

Bibliography: Chapter 2

opinion noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com. https://www.oxfordlearnersdictionaries.com/definition/english/opinion?q=opinion

feedback noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com. https://www.oxfordlearnersdictionaries.com/definition/english/feedback?q=feedback

Agnihotri, A., & Bhattacharya, S. (2016). Online Review Helpfulness: Role of Qualitative Factors: ONLINE REVIEW HELPFULNESS. *Psychology & Marketing*, *33*(11), 1006–1017. https://doi.org/10.1002/mar.20934

Antioco, M., & Coussement, K. (2018). Misreading of consumer dissatisfaction in online product reviews: Writing style as a cause for bias. *International Journal of Information Management*, *38*(1), 301–310. https://doi.org/10.1016/j.ijinfomgt.2017.10.009

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.

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research*, *53*(3), 297–318. https://doi.org/10.1509/jmr.14.0380

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*(2), 99–126.

Barron, G., & Yechiam, E. (2002). Private e-mail requests and the diffusion of responsibility. *Computers in Human Behavior*, *18*(5), 507–520. https://doi.org/10.1016/S0747-5632(02)00007-9

Berkowitz, L. (1972). Social norms, feelings, and other factors affecting helping and altruism. In *Advances in experimental social psychology* (Vol. 6, pp. 63–108). Elsevier.

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, *74*(5), 817–848.

Brewer, M. B., & Kramer, R. M. (1986). *Choice Behavior in Social Dilemmas: Effects of Social Identity, Group Size, and Decision Framing*. 7.

Budescu, D. V., & Chen, E. (2015). Identifying Expertise to Extract the Wisdom of Crowds. *Management Science*, *61*(2), 267–280. https://doi.org/10.1287/mnsc.2014.1909

Carter, C. R., Kaufmann, L., & Michel, A. (2007). Behavioral supply management: A taxonomy of judgment and decision-making biases. *International Journal of Physical Distribution & Logistics Management*.

Chen, P.-Y., Hong, Y., & Liu, Y. (2018a). The value of multidimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Science*, *64*(10), 4629–4647.

Cheung, C. M. K., & Lee, M. K. O. (2012). What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decision Support Systems*, *53*(1), 218–225. https://doi.org/10.1016/j.dss.2012.01.015

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944–957.

Christian, L. M., & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly*, *68*(1), 57–80.

Christian, L. M., Dillman, D. A., & Smyth, J. D. (2007). Helping respondents get it right the first time: The influence of words, symbols, and graphics in web surveys. *Public Opinion Quarterly*, *71*(1), 113–125.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015.
Clarkson, J. J., Tormala, Z. L., Rucker, D. D., & Dugan, R. G. (2013). The malleable influence of social consensus on attitude certainty. *Journal of Experimental Social Psychology*, *49*(6), 1019–1022. https://doi.org/10.1016/j.jesp.2013.07.001
Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4, Pt.1), 377–383. https://doi.org/10.1037/h0025589
de Langhe, B., Fernbach, P. M., & Lichtenstein, D. R. (2016). Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, *42*(6), 817–833. https://doi.org/10.1093/jcr/ucv047
Dillman, D. A., & Christian, L. M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods*, *17*(1), 30–52.
Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? — An empirical investigation of panel data. *Decision Support Systems*, *45*(4), 1007–1016. https://doi.org/10.1016/j.dss.2008.04.001
Epstein, S. (2003). Cognitive-Experiential Self-Theory of Personality. In I. B. Weiner (Ed.), *Handbook of Psychology* (p. wei0507). John Wiley & Sons, Inc. https://doi.org/10.1002/0471264385.wei0507
Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390.
Feldman, J. M., & Lynch, J. G. (1988a). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*(3), 421.
Feldman, J. M., & Lynch, J. G. (1988b). *Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior*. 15.
Fisher, M., Newman, G. E., & Dhar, R. (2018). Seeing stars: How the binary bias distorts the interpretation of customer ratings. *Journal of Consumer Research*, *45*(3), 471–489.
Franzosi, R. (1998). Narrative analysis—Or why (and how) sociologists should be interested in narrative. *Annual Review of Sociology*, *24*(1), 517–554.
Frijda, N. H., & Mesquita, B. (1994). *The social roles and functions of emotions.*
Godek, J., & Murray, K. B. (2008). Willingness to pay for advice: The role of rational and experiential processing. *Organizational Behavior and Human Decision Processes*, *106*(1), 77–87. https://doi.org/10.1016/j.obhdp.2007.10.002
Godes, D., & Silva, J. C. (2012). Sequential and Temporal Dynamics of Online Opinion. *Marketing Science*, *31*(3), 448–473. https://doi.org/10.1287/mksc.1110.0653
Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. *Journal of Consumer Research*, *35*(3), 472–482. https://doi.org/10.1086/586910
Griskevicius, V., Goldstein, N. J., Mortensen, C. R., Cialdini, R. B., & Kenrick, D. T. (2006). Going along versus going alone: When fundamental motives facilitate strategic (non) conformity. *Journal of Personality and Social Psychology*, *91*(2), 281.
Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*. University of Kansas, KS.
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. https://doi.org/10.1002/dir.10073
Herr, P. M., Kardes, F. R., & Kim, J. (1991a). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. *Journal of Consumer Research*, *17*(4), 454–462.
Herr, P. M., Kardes, F. R., & Kim, J. (1991b). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective. *Journal of Consumer Research*, *17*(4), 454–462.

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. https://doi.org/10.1016/j.dss.2017.06.007
Huang, L., Tan, C.-H., Ke, W., & Wei, K.-K. (2013). Comprehension and Assessment of Product Reviews: A Review-Product Congruity Proposition. *Journal of Management Information Systems*, *30*(3), 311–343. https://doi.org/10.2753/MIS0742-1222300311
Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Eighth International AAAI Conference on Weblogs and Social Media*.
Jiménez, F. R., & Mendoza, N. A. (2013). Too Popular to Ignore: The Influence of Online Reviews on Purchase Intentions of Search and Experience Products. *Journal of Interactive Marketing*, *27*(3), 226–235. https://doi.org/10.1016/j.intmar.2013.04.004
Korgaonkar, P., & Wolin, L. D. (2002). Web usage, advertising, and shopping: Relationship patterns. *Internet Research*.
Kovács, B., Carroll, G. R., & Lehman, D. W. (2014). Authenticity and Consumer Value Ratings: Empirical Tests from the Restaurant Domain. *Organization Science*, *25*(2), 458–478. https://doi.org/10.1287/orsc.2013.0843
Kozlov, M. D., & Johansen, M. K. (2010). Real Behavior in Virtual Environments: Psychology Experiments in a Simple Virtual-Reality Paradigm Using Video Games. *Cyberpsychology, Behavior, and Social Networking*, *13*(6), 711–714. https://doi.org/10.1089/cyber.2009.0310
Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, *89*(2), 308.
Li, X., & Hitt, L. M. (2008). Self-Selection and Information Role of Online Product Reviews. *Information Systems Research*, *19*(4), 456–474. https://doi.org/10.1287/isre.1070.0154
Lichtendahl, K. C., Grushka-Cockayne, Y., & Pfeifer, P. E. (2013). The Wisdom of Competitive Crowds. *Operations Research*, *61*(6), 1383–1398. https://doi.org/10.1287/opre.2013.1213
Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates. *Journal of Marketing*, *77*(1), 87–103. https://doi.org/10.1509/jm.11.0560
Lynch Jr, J. G., Marmorstein, H., & Weigold, M. F. (1988). Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. *Journal of Consumer Research*, *15*(2), 169–184.
Makridakis, S., & Winkler, R. L. (1983). Averages of Forecasts: Some Empirical Results. *Management Science*, *29*(9), 987–996. https://doi.org/10.1287/mnsc.29.9.987
Markey, P. M. (2000). Bystander intervention in computer-mediated communication. *Computers in Human Behavior*, *16*(2), 183–188. https://doi.org/10.1016/S0747-5632(99)00056-4
Menon, G., Raghubir, P., & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility-diagnosticity framework. *Journal of Consumer Research*, *22*(2), 212–228.
Moe, W. W., & Schweidel, D. A. (2012). Online product opinions: Incidence, evaluation, and evolution. *Marketing Science*, *31*(3), 372–386.
Moe, W. W., & Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research*, *48*(3), 444–456.
Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon. Com. *MIS Quarterly*, 185–200.
Naylor, R. W., Lamberton, C. P., & Norton, D. A. (2011). Seeing Ourselves in Others: Reviewer Ambiguity, Egocentric Anchoring, and Persuasion. *JouRNal oF MaRketiNg ReSeaRch*, 15.
Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLoS ONE*, *9*(12), e115844. https://doi.org/10.1371/journal.pone.0115844
Petty, R. E., Cacioppo, J. T., Petty, R. E., & Cacioppo, J. T. (1986). *The elaboration likelihood model of persuasion*. Springer.
Petty, R. E., Harkins, S. G., & Williams, K. D. (1980). The effects of group diffusion of cognitive effort on attitudes: An information-processing view. *Journal of Personality and Social Psychology*, *38*(1), 81–92. https://doi.org/10.1037/0022-3514.38.1.81

Polzin, F., Toxopeus, H., & Stam, E. (2018). The wisdom of the crowd in funding: Information heterogeneity and social networks of crowdfunders. *Small Business Economics*, *50*(2), 251–273.
Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, *13*(10), 420–428.
Redline, C., Dillman, D. A., Dajani, A. N., & Scaggs, M. A. (2003). Improving navigational performance in US Census 2000 by altering the visually administered languages of branching instructions. *Journal of Official Statistics*, *19*(4), 403.
Reyna, V. F. (2009). Fuzzy-Trace Theory. In M. Kattan, *Encyclopedia of Medical Decision Making*. SAGE Publications, Inc. https://doi.org/10.4135/9781412971980.n155
Riedl, C., Blohm, I., Leimeister, J. M., & Krcmar, H. (2013). The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. *International Journal of Electronic Commerce*, *17*(3), 7–36. https://doi.org/10.2753/JEC1086-4415170301
Rocklage, M. D., & Fazio, R. H. (2015). The Evaluative Lexicon: Adjective use as a means of assessing and distinguishing attitude valence, extremity, and emotionality. *Journal of Experimental Social Psychology*, *56*, 214–227. https://doi.org/10.1016/j.jesp.2014.10.005
Rocklage, M. D., & Fazio, R. H. (2018). Attitude Accessibility as a Function of Emotionality. *Personality and Social Psychology Bulletin*, *44*(4), 508–520. https://doi.org/10.1177/0146167217743762
Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). Persuasion, Emotion, and Language: The Intent to Persuade Transforms Language via Emotionality. *Psychological Science*, *29*(5), 749–760. https://doi.org/10.1177/0956797617744797
Rutkowski, G. K., Gruder, C. L., & Romer, D. (1983). Group cohesiveness, social norms, and bystander intervention. *Journal of Personality and Social Psychology*, *44*(3), 545.
Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, *311*(5762), 854–856. https://doi.org/10.1126/science.1121066
Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour*, *11*(3), 234–243.
Schneider, C., Weinmann, M., Mohr, P. N., & vom Brocke, J. (2021). When the stars shine too bright: The influence of multidimensional ratings on online consumer ratings. *Management Science*, *67*(6), 3871–3898.
Schoemaker, P. J., & Russo, J. E. (1994). A pyramid of decision approaches. In *Decision theory and decision analysis: Trends and challenges* (pp. 53–78). Springer.
Simon, F., & Usunier, J.-C. (2007). Cognitive, demographic, and situational determinants of service customer preference for personnel-in-contact over self-service technology. *International Journal of Research in Marketing*, *24*(2), 163–173. https://doi.org/10.1016/j.ijresmar.2006.11.004
Surowiecki, J. (2005). The wisdom of crowds. Anchor. (2005).
Tinsley, C. H., Dillon, R. L., & Madsen, P. M. (2011). How to avoid catastrophe. *Harvard Business Review*, *89*(4), 90–97.
Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly*, *68*(3), 368–393.
Tsagkias, M., De Rijke, M., & Weerkamp, W. (2011). *Linking online news and social media*. 565–574.
Tsagkias, M., Weerkamp, W., & Rijke, M. de. (2010). *News comments: Exploring, modeling, and online prediction*. 191–203.
Van de Rijt, A., Kang, S. M., Restivo, M., & Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, *111*(19), 6934–6939.
Van Kleef, G. A., Van den Berg, H., & Heerdink, M. W. (2015). The persuasive power of emotions: Effects of emotional expressions on attitude formation and change. *Journal of Applied Psychology*, *100*(4), 1124.
Watson, J., Ghosh, A. P., & Trusov, M. (2018). Swayed by the Numbers: The Consequences of Displaying Product Review Attributes. *Journal of Marketing*, *82*(6), 109–131. https://doi.org/10.1177/0022242918805468

Wegener, D. T., Petty, R. E., & Klein, D. J. (1994). Effects of mood on high elaboration attitude change: The mediating role of likelihood judgments. *European Journal of Social Psychology*, *24*(1), 25–43.
Weldon, E., & Gargano, G. M. (1985a). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. *Organizational Behavior and Human Decision Processes*, *36*(3), 348–361.
Weldon, E., & Gargano, G. M. (1985b). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. *Organizational Behavior and Human Decision Processes*, *36*(3), 348–361. https://doi.org/10.1016/0749-5978(85)90004-4
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. https://doi.org/10.1509/jmr.13.0379
Yoo, K. H., & Gretzel, U. (2008). What Motivates Consumers to Write Online Travel Reviews? *Information Technology & Tourism*, *10*(4), 283–295. https://doi.org/10.3727/109830508788403114
Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, *37*(2), 197–206. https://doi.org/10.1086/651257
Zhou, S., & Guo, B. (2017). The order effect on online review helpfulness: A social influence perspective. *Decision Support Systems*, *93*, 77–87. https://doi.org/10.1016/j.dss.2016.09.016

Bibliography: Chapter 3

Ahmad, Shimi N. and Michael Laroche (2017), "Analyzing Electronic Word of Mouth: A Social Commerce Construct". *International Journal of Information Management*, 37 (3), 202-213.

Alba, Joseph W., and J. Wesley Hutchinson (1987), "Dimensions of consumer expertise." *Journal of Consumer Research*, 13 (4), 411-454.

Babin, Laurie A., and Alvin C. Burns (1997), "Effects of print ad pictures and copy containing instructions to imagine on mental imagery that mediates attitudes." *Journal of Advertising*, 26 (3), 33-44.

Baker, Andrew M., Naveen Donthu, and Vineet Kumar (2016), "Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions." *Journal of Marketing Research*, 53 (2), 225-239.

Barasch, Alixandra, Kristin Diehl, Jackie Silverman, and Gal Zauberman, (2017), "Photographic memory: The effects of volitional photo taking on memory for visual and auditory aspects of an experience." *Psychological Science*, 28 (8), 1056-1066.

Barasch, Alixandra, Gal Zauberman, and Kristin Diehl (2018), "How the intention to share can undermine enjoyment: Photo-taking goals and evaluation of experiences." *Journal of Consumer Research*, 44 (6), 1220-1237.

Bell, Brad E., and Elizabeth F. Loftus (1989), "Trivial persuasion in the courtroom: The power of (a few) minor details." *Journal of Personality and Social Psychology*, 56 (5), 669-679.

Bigne, Enrique., Kalliopi Chatzipanagiotou, and Carla Ruiz (2020), "Pictorial content, sequence of conflicting online reviews and consumer decision-making: The stimulus-organism-response model revisited." *Journal of Business Research*, 115, 403-416.

Brinol, Pablo and Richard E. Petty (2009), "Source factors in persuasion: A self-validation approach.' *European Review of Social Psychology*, 20 (1), 49-96.

Browne, Glenn, Mitzi G. Pitts, and James C. Wetherbe (2007), "Cognitive stopping rules for terminating information search in online tasks". *MIS Quarterly*, 31 (1), 89-104.

Boulding, William, and Amna Kirmani (1993), "A consumer-side experimental examination of signaling theory: do consumers perceive warranties as signals of quality?." *Journal of Consumer Research*, 20 (1), 111-123.

Cao, Qing, Wenjing Duan, and Qiwei Gan (2011), "Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach". *Decision Support Systems*, 50 (2), 511-521.

Cardwell, Brittany A., Linda A. Henkel, Maryanne Garry, Eryn J. Newman, and Jeffrey L. Foster (2016), "Nonprobative photos rapidly lead people to believe claims about their own (and other people's) pasts." *Memory & Cognition*, 44 (6), 883-896.

Cardwell, Brittany A., Eryn J Newman, Maryanne Garry, Antonia Mantonakis, and Randi Beckett (2017), "Photos that increase feelings of learning promote positive evaluations". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43 (6), 944 - 954.

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

De Langhe, Bart, Philip M. Fernbach, and Donald R. Lichtenstein (2016), "Navigating by the stars: Investigating the actual and perceived validity of online user ratings." Journal of Consumer Research, 42 (6), 817-833.

Derksen, Daniel G., Megan E. Giroux, Deborah A. Connolly, Eryn J. Newman, and Daniel M. Bernstein (2020), "Truthiness and law: Nonprobative photos bias perceived credibility in forensic contexts". *Applied Cognitive Psychology*, 34 (6), 1335-1344.

Diehl, Kristin, and Gal Zauberman (2022), "Capturing life or missing it: how mindful photo-taking can affect experiences." *Current Opinion in Psychology*, 46.

Diehl, Kristin, Gal Zauberman, and Alixandra Barasch (2016), "How taking photos increases enjoyment of Experiences". *Journal of Personality and Social Psychology*, 111 (2), 119-140.

Dunaway, Johanna, Kathleen Searles, Mingxiao Sui, and Newly Paul (2018), "News attention in

a mobile era." *Journal of Computer-Mediated Communication*, 23 (2), 107-124.
Edell, Julie A., and Richard Staelin (1983), "The information processing of pictures in print advertisements." *Journal of Consumer Research*, 10 (1), 45-61.
Fenn, Elise, Eryn J. Newman, Kathy Pezdek, and Maryanne Garry (2013), "The effect of nonprobative photographs on truthiness persists over time". *Acta Psychologica*, 144 (1), 207-211.
Fenn, Elise, Nicholas Ramsay, Justin Kantner, Kathy Pezdek, and Erica Abed (2019), "Nonprobative photos increase truth, like, and share judgments in a simulated social media environment". *Journal of Applied Research in Memory and Cognition*, 8 (2), 131-138.
Filieri, Raffaele (2015), "What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM". *Journal of Business Research*, 68 (6), 1261-1270.
______ (2016), "What makes an online consumer review trustworthy?" *Annals of Tourism Research,* 58, 46-64.
He, Stephen X., and Samuel D. Bond (2015), "Why is the crowd divided? Attribution for dispersion in online word of mouth." *Journal of Consumer Research*, 41 (6), 1509-1527.
Henkel, Linda A. (2014), "Point-and-shoot memories: The influence of taking photos on memory for a museum tour". *Psychological Science*, 25 (2), 396-402.
Hertzendorf, Mark N. (1993) "I'm not a High-Quality Firm-but I Play One on TV". *The RAND Journal of Economics*, 24 (2), 236-247.
Hlee, Sunyoung, Jimin Lee, Sung-Byung Yang, and Chulmo Koo (2019), "The moderating effect of restaurant type on hedonic versus utilitarian review evaluations." International Journal of Hospitality Management, 77, 195-206.
Jalilvand, Mohammad Reza, and Neda Samiei (2012), "The effect of electronic word of mouth on brand image and purchase intention: An empirical study in the automobile industry in Iran." *Marketing Intelligence & Planning*, 30 (4), 460-476.
Kats, Rimma (2001), "User-generated visual content can influence purchases," *eMarketer*, February 21. https://www.emarketer.com/content/user-generated-visual-content-influence-purchases
Kihlstrom, Richard E., and Michael H. Riordan (1984), "Advertising as a Signal. *Journal of Political Economy"*, 92 (3), 427-450.

Kronrod, Ann, and Shai Danziger (2013), ""Wii will rock you!" The use and effect of figurative language in consumer reviews of hedonic and utilitarian consumption." *Journal of Consumer Research*, 40 (4), 726-739.
Li, Xin, Rob Law, Gang Xie, and Shouyang Wang (2021) "Review of tourism forecasting research with internet data." *Tourism Management*, 83.
Lu, Shuya, Jianan Wu, and Shih-Lun Allen Tseng (2018), "How online reviews become helpful: A dynamic Perspective". *Journal of Interactive Marketing*, 44, 17-28.
Ma, Yufeng, Zheng Xiang, Qianzhou Du, and Weiguo Fan (2018), "Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning". *International Journal of Hospitality Management*, 71, 120-131.
Mafael, Alexander (2019), "How regulatory orientation and feelings of gratitude shape online review Helpfulness". *Journal of Consumer Psychology*, 29 (4), 601-622.
Mantonakis, Antonia, Brittany Cardwell, Randi Beckett, Eryn Newman, and Maryanne Garry (2014), "The mere presence of a photo on a product label can change taste perception." In *Conference Proceedings of the 8th International Conference of the Academy of Wine Business Research*, Hochschule Geisenheim University, Germany.
Marder, Ben, Antonia Erz, Rob Angell, and Kirk Plangger (2021), "The role of photograph aesthetics on online review sites: Effects of management-versus traveler-generated photos on tourists' decision making". *Journal of Travel Research*, 60 (1), 31-46.
Milgrom, Paul, and John Roberts (1986), "Price and advertising signals of product quality". *Journal of Political Economy,* 94 (4), 796-821.
Miniard, Paul W., Sunil Bhatla, Kenneth R. Lord, Peter R. Dickson, and H. Rao Unnava (1991),

"Picture-based persuasion processes and the moderating role of involvement." *Journal of Consumer Research*, 18 (1), 92-107.
Mitchell, Andrew A., and Jerry C. Olson (1981), "Are product attribute beliefs the only mediator of advertising effects on brand attitude?" *Journal of Marketing Research*, 18 (3), 318-332.
Moore, Sarah G., and Katherine C. Lafreniere (2020), "How online word-of-mouth impacts r eceivers." *Consumer Psychology Review*, I3 (1), 34-59.
Mudambi, Susan M., and David Schuff (2010), "Research note: What makes a helpful online review? A study of customer reviews on Amazon. com." *MIS Quarterly*, 185-200.
Nelson, Phillip (1974)," Advertising as information". *Journal of Political Economy*, 82 (4), 729-754.
Newman, Eryn J., Tanjeem Azad, D. Stephen Lindsay, and Maryanne Garry (2018), "Evidence that photos promote rosiness for claims about the future." *Memory & Cognition*, 46 (8), 1223-1233.
_________., Maryanne Garry, Daniel M. Bernstein, Justin Kantner, and D. Stephen Lindsay (2012), "Nonprobative photographs (or words) inflate truthiness". *Psychonomic Bulletin & Review*, 19 (5), 969-974.
_________., Maryanne Garry, Christian Unkelbach, Daniel M. Bernstein, D. Stephen Lindsay, and Robert A. Nash (2015), "Truthiness and falsiness of trivia claims depend on judgmental contexts". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41 (5), 1337.

_________., and Lynn Zhang (2020). Truthiness: How non-probative photos shape belief. in *The Psychology of Fake News*, ed. Rainer Greifeneder, Mariela Jaffe, Eryn Newman, and Norbert Schwarz, Routledge, (pp. 90-114).
Nightingale, Sophie J., Kimberley A. Wade, and Derrick G. Watson (2017), "Can people identify original and manipulated photos of real-world scenes?" *Cognitive Research: Principles and Implications,* 2 (1), 1-21.
Park, Cheol Woo, Ian Sutherland, and Seul Ki Lee (2021), "Effects of online reviews, trust, and picture-superiority on intention to purchase restaurant services." Journal of Hospitality and Tourism Management, 47, 228-236.
Pennebaker, James W., Martha E. Francis, and Roger J. Booth (2001), "Linguistic inquiry and word count: LIWC 2001". *Mahway: Lawrence Erlbaum Associates*, 71 (2001), 2001.
Peracchio, Laura A., and Joan Meyers-Levy (1994), "How ambiguous cropped objects in ad photos can affect product evaluations." *Journal of Consumer Research*, 21 (1), 190-204.
_________ (2005), "Using stylistic properties of ad pictures to communicate with consumers". *Journal of Consumer Research*, 32 (1), 29-40.
Petty, Richard E., Pablo Brinol, and Zakary L. Tormala (2002), "Thought confidence as a determinant of persuasion: the self-validation hypothesis". *Journal of Personality and Social Psychology*, 82 (5), 722.
Ridgeway, Cecilia L., Joseph Berger, and LeRoy Smith (1985), "Nonverbal cues and status: An expectation states approach." *American Journal of Sociology*, 90 (5), 955-978.
Rocklage, Matthew D., and Russell H. Fazio (2020), "The enhancing versus backfiring effects of positive emotion in consumer reviews." *Journal of Marketing Research*, 57 (2), 332-352.

Schellekens, Gaby AC, Peeter WJ Verlegh, and Ale Smidts (2010), "Language abstraction in word of mouth." *Journal of Consumer Research*, 37 (2), 207-223.
Schlosser, Anne E. (2011), "Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments". *Journal of Consumer Psychology*, 21 (3), 226-239.
Shihab, Muhammad Rifki, and Audry Pragita Putri (2019), "Negative online reviews of popular products: Understanding the effects of review proportion and quality on consumers' attitude and intention to buy." *Electronic Commerce Research*, 19 (1), 159-187.
Smith, Ruth Ann (1991), "The effects of visual and verbal advertising information on consumers' inferences." Journal of Advertising, 20 (4), 13-24.
Sims, Lauren (2020), "Product review photos are here," *Trustpilot*, May 18.

https://www.trustpilot.com/blog/reviews-matter/product-review-photos-are-here
Siering, Michael, Jan Muntermann, and Balaji Rajagopalan (2018), “Explaining and predicting online review helpfulness: The role of content and reviewer-related signals”. *Decision Support Systems*, 108, 1-12.
Spence, Michael (2002), "Signaling in retrospect and the informational structure of markets." *American Economic Review*, 92 (3), 434-459.
Street Jr, Richard L., Robert M. Brady, and Raymond Lee (1984), "Evaluative responses to communicators: The effects of speech rate, sex, and interaction context”. *Western Journal of Speed Communication*, 48 (1), 14-27.
Sundar, Shyam S. (2008), *The MAIN model: A heuristic approach to understanding technology effects on credibility,* Cambridge, MA; MacArthur Foundation Digital Media and Learning Initiative.
Taylor, David G. (2020), “Putting the “self” in selfies: how narcissism, envy and self-promotion motivate sharing of travel photos through social media”. *Journal of Travel & Tourism Marketing*, 37 (1), 64-77.
Togawa, Taku, and Yoko Sugitani (2021), “Looks Far Beyond My Reach: The Zoom Effect in Product Videos Influences Luxury Perception and Purchase Intention”. *Journal of Consumer Psychology,* Advance Online Publication. https://doi.org/10.1002/jcpy.1268.
Tonietto, Gabriela N., and Alixandra Barasch (2021), "Generating content increases enjoyment by immersing consumers and accelerating perceived time." *Journal of Marketing*, 85 (6), 83-100.
Tormala, Zakary L. (2016), "The role of certainty (and uncertainty) in attitudes and persuasion." *Current Opinion in Psychology*, 10, 6-11.
Tsai, Claire I., Joshua Klayman, and Reid Hastie (2008), "Effects of amount of information on judgment accuracy and confidence." *Organizational Behavior and Human Decision Processes*, 107 (2), 97-105.
Vana, Prasad, and Anja Lambrecht (2021), "The effect of individual online reviews on purchase likelihood." *Marketing Science*, 40 (4), 708-730.
Watson, Jared, Anastasiya Pocheptsova Ghosh, and Michael Trusov (2018), "Swayed by the numbers: the consequences of displaying product review attributes." *Journal of Marketing*, 82 (6), 109-131.
Wilson, Joseph C., and Deanne L. Westerman (2018), "Picture (im) perfect: Illusions of recognition memory produced by photographs at test." *Memory & Cognition*, 46, (7), 1210-1221.
Xia, Huosong, Xiaoting Pan, Yanjun Zhou, and Zuopeng Justin Zhang (2020), "Creating the best first impression: Designing online product photos to increase sales." *Decision Support Systems*, 131.
Yang, Sung-Byung, Seung-Hun Shin, Youhee Joun, and Chulmo Koo (2017), “Exploring the comparative importance of online hotel reviews’ heuristic attributes in review helpfulness: a conjoint analysis approach." *Journal of Travel & Tourism Marketing*, 37 (7), 963-985.
Zhang, Lynn, Eryn J. Newman, and Norbert Schwarz (2021), “When photos backfire: Truthiness and falsiness effects in comparative judgments”. *Journal of Experimental Social Psychology*, 92, 104054.
Zhang, Kem Z.K., Sesia J. Zhao, Christy M.K Cheung, and Matthew K.O Lee (2014), "Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model." *Decision Support Systems*, 67, 78-89.
Zinko, Robert, Helene de Burgh-Woodman, Zhan Zhang Furner, and Soo Jung Kim (2021), "Seeing is believing: the effects of images on trust and purchase intent in eWOM for hedonic and utilitarian products." *Journal of Organizational and End User Computing*, 33 (2), 85-104.

Bibliography: Chapter 4

Alcañiz, M., Bigné, E., & Guixeres, J. (2019). Virtual reality in marketing: A framework, review, and research agenda. *Frontiers in Psychology*, *10*, 1530.

Brauer, M., Judd, C. M., & Gliner, M. D. (1995). The effects of repeated expressions on attitude polarization during group discussions. *Journal of Personality and Social Psychology*, *68*(6), 1014–1029. https://doi.org/10.1037/0022-3514.68.6.1014

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482.

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. https://doi.org/10.2753/JEC1086-4415130402

Eisingerich, A. B., Auh, S., & Merlo, O. (2014). Acta non verba? The role of customer participation and word of mouth in the relationship between service firms' customer satisfaction and sales performance. *Journal of Service Research*, *17*(1), 40–53.

Flavián, C., Akdim, K., & Casaló, L. V. (2023). Effects of voice assistant recommendations on consumer behavior. *Psychology & Marketing*, *40*(2), 328–346.

Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, *100*, 547–560.

Janis, I. L. (1983). *Groupthink*. Houghton Mifflin Boston.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343.

Leibenstein, H. (1950). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics*, *64*(2), 183. https://doi.org/10.2307/1882692

Shteynberg, G., Hirsh, J. B., Galinsky, A. D., & Knight, A. P. (2014). Shared attention increases mood infusion. *Journal of Experimental Psychology: General*, *143*(1), 123–130. https://doi.org/10.1037/a0031549

Yang, S., Zhou, C., & Chen, Y. (2021). Do topic consistency and linguistic style similarity affect online review helpfulness? An elaboration likelihood model perspective. *Information Processing & Management*, *58*(3), 102521.

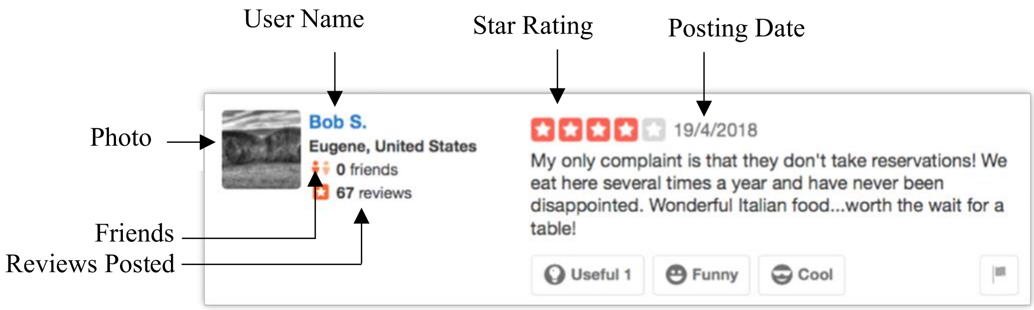
Zarantonello, L., & Schmitt, B. H. (2023). Experiential AR/VR: a consumer and service framework and research agenda. *Journal of Service Management*, *34*(1), 34–55.

Zarouali, B., Brosius, A., Helberger, N., & De Vreese, C. H. (2021). WhatsApp marketing: A study on WhatsApp brand communication and the role of trust in self-disclosure. *International Journal of Communication*, *15*, 25.

Appendixes

Appendix A Study 1a and Study 1b:

Study 1A – Sample Yelp Review



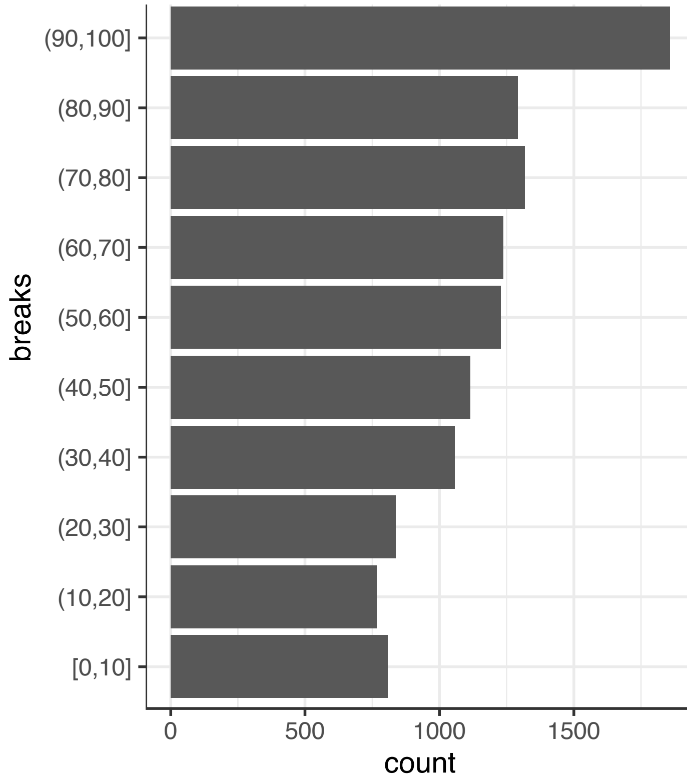
Examples of opinions Analytical scores

Evaluation	Analytical
Excellent place for breakfast or brunch. The burrito is delicious. A variety of great menu options. Traditional breakfast items along with unique options. Around 10 dollars a plate.	99
Came here on a recommendation from a co-worker that attended college near here. In downtown Columbia, in the middle of several college campuses.	99
Before you board Amtrak stop in for a fantastic breakfast. Healthy tasty menu at a reasonable price with friendly service.	99
I had heard a lot about this place so I was pretty excited. Yelp, and my real-life source both agreed on the superiority of their tomato soup, and since I had become a recent tomato soup convert, it seemed the stars had aligned.	50
Shotgun Pete's is the best BBQ in the area. I love their pulled pork sandwich. It is a little pricey which is the main reason I don't visit more often, but I'd definitely recommend it for anyone looking for some great BBQ!	40
Me and my sister stopped in the other day for lunch as we had been craving some good BBQ...and boy were we blown away!	40
Love when places live up to the yelp hype! Its pretty cheap and everything I had was fresh and left me wanting more even though I was stuffed.	5
Salads here are super delicious and a great deal. I also really love their piña coladas. Service has always been good.	5

Examples of sentences VADER sentiment analysis

Sentence	Positive Index	Negative Index	Neutral Index	Compound
VADER is smart, handsome, and funny.	0,75	0,00	0,25	0,83
VADER is smart, handsome, and funny!	0,75	0,00	0,25	0,84
VADER is very smart, handsome, and funny.	0,70	0,00	0,30	0,85
VADER is VERY SMART, handsome, and FUNNY.	0,75	0,00	0,25	0,92
VADER is VERY SMART, handsome, and FUNNY!!!	0,77	0,00	0,23	0,93
VADER is VERY SMART, uber handsome, and FRIGGIN FUNNY!!!	0,71	0,00	0,29	0,95
VADER is not smart, handsome, nor funny.	0,00	0,65	0,35	0,74
The book was good.	0,49	0,00	0,51	0,44
At least it isn't a horrible book.	0,36	0,00	0,64	0,43
The book was only kind of good.	0,30	0,00	0,70	0,38
The plot was good, but the characters are uncompelling and the dialog is not great.	0,09	0,33	0,58	0,70
Today SUX!	0,00	0,78	0,22	0,55
Today only kinda sux! But I'll get by, lol	0,32	0,13	0,56	0,52
Make sure you :) or :D today!	0,71	0,00	0,29	0,86
Catch utf8 emoji such as 💘 and 💋 and 😁	0,28	0,00	0,72	0,70
Not bad at all	0,49	0,00	0,51	0,43

Frequency of Analytic score in intervals of 10



Correlation table for Restaurant dataset

	NthReview	*Rating*	*Elapsed Days*	*Word Count*	*Restaurant Average*	*City*	*Friends*	*Reviews*
NthReview		0.051***	0.583***	0.007	0.043***	0.333***	-0.016	-0.111***
Rating	0.051***		0.032***	-0.153***	0.264***	-0.068***	0.010	-0.013
Elapsed Days	0.583***	0.032***		0.008	0.012	-0.044***	-0.012	-0.087***
Word Count	0.007	-0.153***	0.008		-0.031***	0.011	0.002	-0.004
Restaurant Average	0.043***	0.264***	0.012	-0.031***		-0.235***	-0.016	-0.015
City	0.333***	-0.068***	-0.044***	0.011	-0.235***		-0.004	-0.004
Friends	-0.016	0.010	-0.012	0.002	-0.016	-0.004		0.548***
Reviews	-0.111***	-0.013	-0.087***	-0.004	-0.015	-0.004	0.548***	

Computed correlation used pearson-method with listwise-deletion.

Correlation table for Hotel dataset

	NthReview	*Rating*	*Word Count*	*Restaurant Average*
NthReview		-0.014	-0.012	0.018
Rating	-0.014		-0.127^{***}	0.581^{***}
Word Count	-0.012	-0.127^{***}		-0.023^{*}
Restaurant Average	0.018	0.581^{***}	-0.023^{*}	

Computed correlation used pearson-method with listwise-deletion.

Table of VIF- Restaurant dataset

Variable	*VIF regression model-Compound*	*VIF regression model-Analytic*
LogNtheReview	6.36	4.56
Dummy positive reviews	22.97	19.84
Log Interaction Term	27.77	22.80
Word count	1.03	1.02
Elapsed days	1.80	1.65
Restaurant Average	1.13	1.06
City	1.39	1.43
Friends	1.38	1.47
Reviews	1.42	4.56

Table of VIF- Restaurant dataset

	VIF regression model-Compound	*VIF regression model-Analytic*
LogNtheReview	3.68	1.00
Dummy positive reviews	19.70	
Word Count	1.02	1.00
Restaurant Average	1.33	1.00
Log Interaction Term	22.13	

Alternative models for robustness:

$$Analytic_{ijk} = \alpha + \beta_1 NthReview_i + \beta_2 NthReview_i^2 + \beta_3 Rating_k + \beta_4 Elapseddays_k + \beta_5 WordCount + \beta_6 Restaurant\ Average_j + \beta_7 City_j + \beta_8 Friends_i + \beta_9 Reviews_i + \varepsilon_{ijk}$$

$$Compound_{ijk} = \alpha + \beta_1 NthReview_i + \beta_2 NthReview_i^2 + \beta_3 Rating_k + \beta_4 Elapseddays_k + \beta_5 WordCount + \beta_6 Restaurant\ Average_j + \beta_7 City_j + \beta_8 Friends_i + \beta_9 Reviews_i + \varepsilon_{ijk}$$

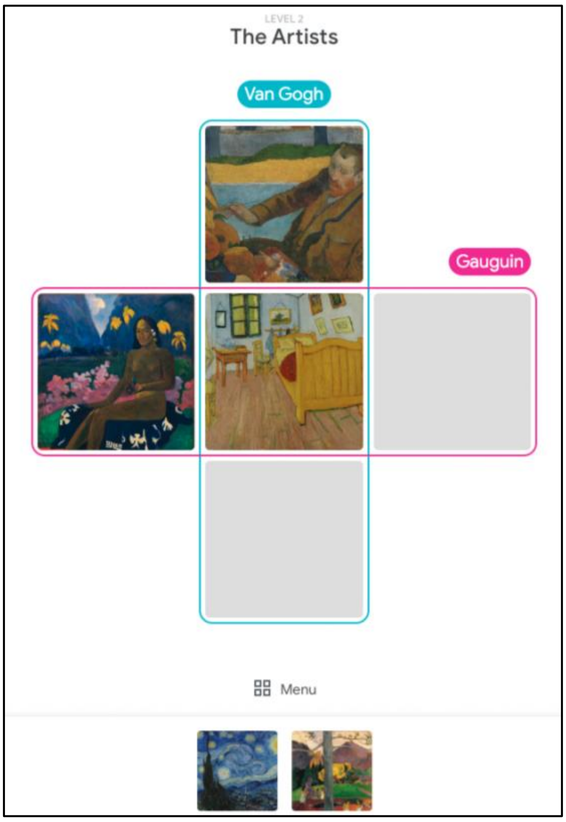
$$Tone = \alpha + \beta_1 LogNthReview_i + \beta_2 DummyPositiveReviews_k + \beta_3 LogInteraction\ term_{ik} + \beta_4 WordCount_k + \beta_5 Elapseddays_k + \beta_6 Restaurant\ Average_j + \beta_7 Friends_i + \beta_8 Reviews_i + \varepsilon_{ijk}$$

Results of Alternative robustness models

	Analytic	Compound	Tone
(Intercept)	66.958 ***	-0.003	58.792 ***
	(3.002)	(0.041)	(4.006)
NthReview	-0.013 ***	0.000	
	(0.003)	(0.000)	
Dummy positive reviews	4.868 ***	0.658 ***	18.858 ***
	(0.772)	(0.010)	(3.301)
Word Count	-0.027 ***	0.002 ***	0.017 **
	(0.006)	(0.000)	(0.006)
Elapsed days	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Restaurant Average	-2.723 ***	0.001	0.126
	(0.757)	(0.010)	(0.717)
Friends	-0.002	-0.000	-0.002
	(0.001)	(0.000)	(0.001)
Reviews	0.006 ***	-0.000 ***	-0.008 ***
	(0.002)	(0.000)	(0.002)
Log Nth Review			-1.541 *
			(0.676)
Log Interaction Term			3.589 ***
			(0.696)
AIC	96221.649	9289.869	95110.875
BIC	96286.635	9354.855	95183.081
Log Likelihood	-48101.825	-4635.935	-47545.437
Deviance	8073943.153	1480.971	7231958.639
Num. obs.	10104	10104	10104

Appendix B- Study 2 Manipulating "N" For a Review of Visual Crosswords

Visual Crossword Game



Allocated 100 points to "informational aspect" versus "personal experience":

- What would you include in your review of this visual crossword so that your review adds value?

Scale:

- Informational Aspects [0-100]
- Personal Experience [0-100]

Baseline Responsibility measure:

- In general, how much responsibility do you think the 1st reviewer would feel in posting their opinion?

Scale:
1 = "No responsibility", 2 = "A little responsibility", 3 = "Some responsibility", 4 = "A lot of responsibility", 5 = "All responsibility"

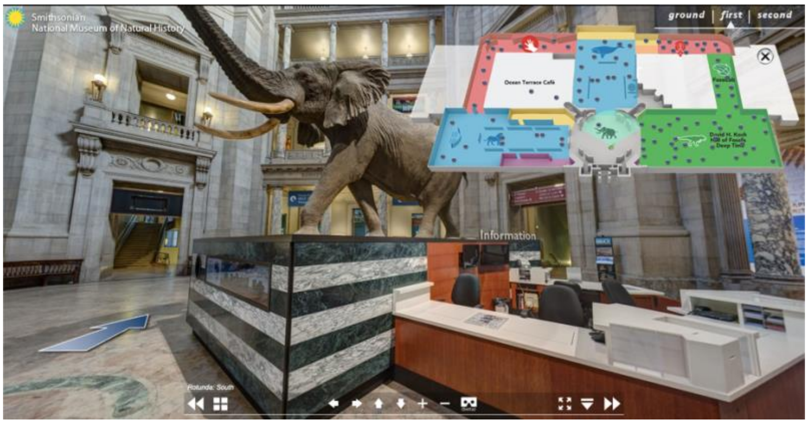
Responsibility measure:

- How responsible do you feel to provide an informative review?
- How responsible do you feel to provide a carefully thought-out review?
- How responsible do you feel to provide a helpful review?
- How responsible do you feel to provide a detailed review?

Scale:
1 = "Much less", 2 = "A moderate amount less", 3 = "Slightly less," 4 = "About the same", 5 = "Slightly more", 6 = "Moderately more", 7 = "Much more"

Appendix C- Study 3 Role of social norms

Virtual tour



How participants would edit review text:

- Adding more detailed description of the virtual tour vs. Adding more emotional impact to your writing.
- Providing information about the virtual tour vs. Giving more of a sense of how you felt.

Scale:

- Adding more detailed description of the virtual tour [0-100]
- Adding more emotional impact to your writing [0-100]

&

- Providing information about the virtual tour [0-100]
- Giving more of a sense of how you felt [0-100]

Appendix D- Study 4 Moderating Role of Feedback vs Opinion

Allocated 100 points to “informational aspect” versus “personal experience”:

- What would you include in your review of this visual crossword so that your review adds value?

Scale:

- Informational Aspects [0-100]
- Personal Experience [0-100]

Appendix E- Study 5 Single vs Multi Dimensional Ratings

Single dimensional rating structure:

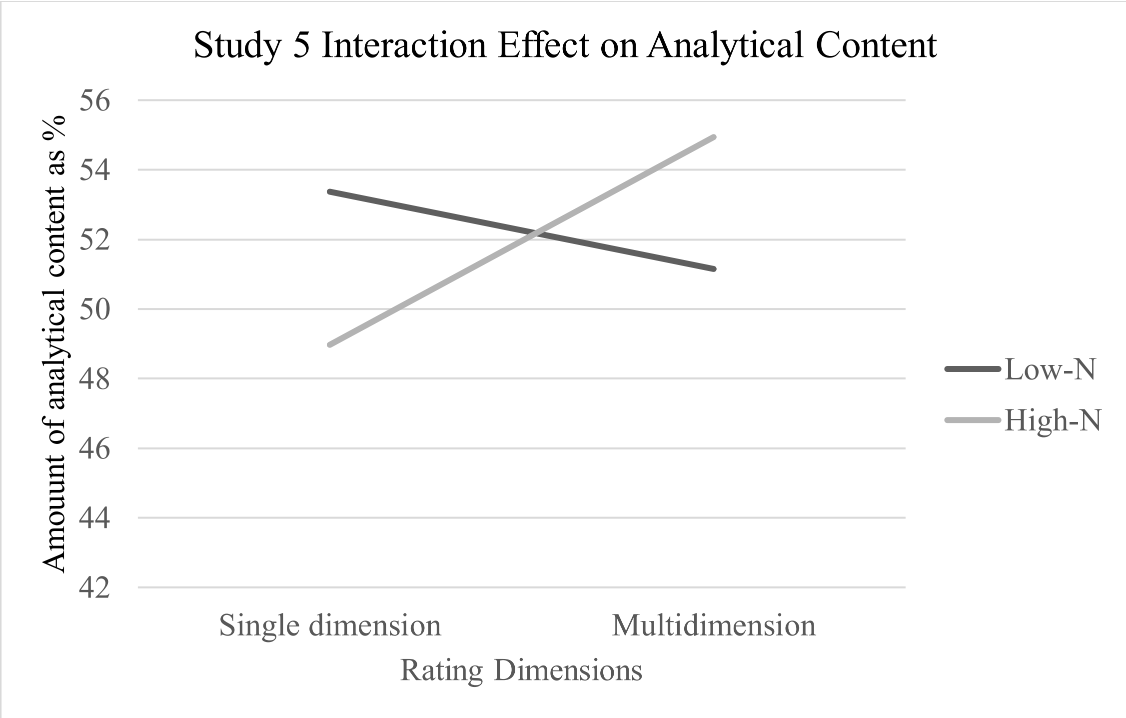
- Overall

Multidimensional rating structure

- Overall
- Image quality
- Ease of navigation
- Collection of exhibits
- Layout of virtual tour webpage

Scale:

- 1 to 5 stars with half stars.



Appendix F: TripAdvisor Photos in Study 2

Task instructions: *In this study, you will see photos of London landmarks posted by real people. Some of these photos are from positive reviews (4 or 5 stars) and some are from negative reviews (1 or 2 stars). You will also see two review texts below each photo. Your task will be to guess which of the texts was uploaded along with the photo you see.*

Sample trial: *"The photos below are all from the Shard Building in London. Please indicate which of the reviews you think was uploaded with the photo."*

Scale: Binary choice between the two review texts.

Photos from TripAdvisor on views from the Shard	Sample reviews (Note: Review A was the correct review…)
	**Review A:** "It's a magnificent building, just to stare at, but the fast elevators whisk you up to wonderful views of the city, the river even The London Eye. Plan on spending some time as there are high top cocktail tables with chairs, snacks are available and most drinks."
	Review B: "You can't beat this view. Don't miss this view of London. We almost passed it by. So, glad we changed our mind. It is a little pricey if you buy a single ticket, but ours was included in the London Pass, so it was definitely worth the stop"
	**Review A:** "Once in a lifetime. Fantastic experience, went up at night time and the view was amazing. Didn't take as long as I thought it would though."
	Review B: "Just a simple 'No' for me. In my opinion, the Shard is OK, it provide a great view of London. But the problem is, it is so overly priced, the bar at the top is very expensive and to be honest, it is just a very tall tower with a view of the city. Nothing really special."

Appendix G: Pizza Scenario Stimuli in Study 3

Task instructions: *In this study, you will see a sample of online posts from customers who have visited a pizza restaurant. You will see a star rating and a review. Both posts have a positive [negative] rating. While looking at these posts, please imagine that you are thinking about going to the restaurant in real life and that you are looking at online reviews.*

Look carefully at each post. You will be asked to indicate how helpful and useful the customers were with their posts.

Review valence:	Probative text:	Nonprobative text:	Photo
Positive (4 stars out of 5)	"Good pizza. Balance of fresh, healthy ingredients and creative recipes."	"Delicious food and friendly staff. Good atmosphere too."	
Negative (2 stars out of 5)	"The pizza fell apart when you picked it up, and it was messy."	"Food wasn't great...low quality ingredients. Staff were slow."	

Dependent measure:
- Which customer's post, A or B, was relatively more helpful?
- Which customer's post, A or B, was relatively more useful?

1 = "Much more Post A", 2 = "More Post A", 3 = "Slightly more Post A," 4 = "About the same", 5 = "Slightly more Post B", 6 = "More Post B", 7 = "Much more Post B"

Pre-test for Study 3 stimuli:

$N = 80$, 50% female, $M_{Age} = 39.21$, $SD = 10.91$ on Prolific:

Positive photo and reviews stimuli:

Photo	**Pre- test study 3**
	• **Customer A:** "Delicious food and friendly staff. The atmosphere was good too..” • **Customer B:** “Good pizza. Balance of fresh, healthy ingredients and creative recipes."

Negative photo and reviews stimuli:

	• **Customer C:** "Food wasn't great...low quality ingredients. Staff were slow. " • **Customer D:** “The pizza fell apart when you picked it up, and it was messy.”

Participants were asked to look at a photo with two reviews and answer the questions that follow.

- *Positive set*: In your opinion, which of the two customers' reviews, A or B, is better illustrated by the photo above? (96.25% chose Customer B)
- *Negative set*: In your opinion, which of the two customers' reviews, C or D, is better illustrated by the photo above? (97.5 % chose Customer D)

For each review, participants were also asked to rate their agreement with two statements about the probative value of the review on a 5-point scale (1 = “strongly disagree”, 5 = “strongly agree):

(1) “Claims in this customer’s review could be evaluated by looking at this photo.”
(2) “The photo depicts elements of this customer’s review.”

The two items were combined for an index of how probative the review was, in relation to the attached photo. See means and standard deviations below:

	Nonprobative reviews (A & C)		Probative reviews (B & D)	
Review valence:	*M*	*SD*	*M*	*SD*
Positive (A & B)	3.02	0.84	4.61	0.49
Negative (C & D)	2.87	0.94	4.53	0.61

Further, in the A and C reviews, the mean agreement was not significantly different from the scale mid-point, indicating little agreement that these reviews were probative, in relation to the attached photo. Within each valence, agreement with these statements was significantly greater in the probative (vs. nonprobative) conditions (both $t(80)>14.97$, $p < .002$).

Appendix H: BBQ pit stimuli in Study 4

Marketer-provided information on the BBQ pits:
•Portable: Enjoy the best flavor of an outdoor barbecue any place and any time.
•Durable steel: Cook in a traditional way without burning or damaging from frequent use.
•Easy clean: Wipe with a damp cloth to remove the residue.
•Ready to use: Simply put the folding legs in to start barbecuing.
•Easy to assemble: removable grill and mess-free ash catcher for easy clean-up.
•Smooth ventilation: air flows well and charcoal burns efficiently, so your food cooks faster.

BBQ A	BBQ B
	

Photos included with the reviews:

Study 4 review text, ratings of helpfulness and valence in the pre-test (*M*=58 on mTurk):

		Helpfulness (1 to 5)		Valence (-5 to +5)	
Set	Review text	*M*	*SD*	*M*	*SD*
Positive	"Works well. Happy with this purchase."	2.98	1.00	3.26	1.84
	"It works. Have used it constantly where weather has allowed."	3.22	1.19	2.69	2.19
	"Assemble was intuitive, easy to use. Looks good in the garden."	3.84	0.91	3.38	1.54
	"Used this for our first BBQ of the season and I very was impressed."	3.36	1.02	3.64	1.73
	"A fantastic bbq. Reasonably simple to put together, looks great and does what you want."	3.53	0.96	3.48	1.70
	"Have used it frequently. Okay for the price."	3.03	0.95	1.66	1.67
Negative/ Neutral	"An alright BBQ...lacks some stability."	3.88	0.92	-2.19	2.90
	"Not the best quality BBQ. Had temperature issues..."	3.71	0.88	-1.16	2.67
	"Found it difficult to keep hot enough to cook anything quickly."	3.53	0.99	-0.34	2.57
	"Had a few issues with stability and cleaning."	3.71	0.97	-1.34	3.29
	"Great that it folds flat. But rusted after one wash."	3.76	0.94	-0.12	2.56
	"Nice design and sturdy but It rusted after one use."	3.67	0.85	0.02	2.74

Measures in Study 4a:

Relative choice likelihood:

- Suppose that you only had the posts above to make your decision. How likely would you be to pick BBQ A vs BBQ B?

Scale:
1 = "Much more likely A", 2 = "More likely A", 3 = "Slightly more likely A," 4 = "About the same", 5 = "Slightly more likely B", 6 = "More likely B", 7 = "Much more likely B"

Helpfulness:

- How helpful were the posts of BBQ A relative to BBQ B? (No photo condition)
- How helpful were the posts with photos in making your decision between BBQ A and B? (Photo condition)

Scale:
1 = "Much less helpful", 2 = "Less helpful", 3 = "Moderately less helpful," 4 = "About the same", 5 = "Moderately more helpful", 6 = "More helpful", 7 = "Much more helpful"

(Note: Results of the *Helpfulness* measure were significant and in the direction expected, but these results were not included in the manuscript for the purpose of brevity.)

Confidence:

- For which BBQ pit are you more confident that it is popular with customers?
- For which BBQ pit are you more confident that customers would recommend it to others?

- For which BBQ pit are you more confident that customers are satisfied with it?

Scale:
1 = "A lot more confident in A", 2 = "More confident in A", 3 = "Slightly more confident in A," 4 = "About the same", 5 = "Slightly more confident in B", 6 = "More confident in B", 7 = "A lot more confident in B"

Measures in Study 4b:

Relative choice likelihood:

- Suppose that you only had the posts above to evaluate these two BBQ pits. As you are narrowing down your options, you will reject some options and look into other options further. If you were screening out options and had to drop one of the BBQ pits, which one would you be more likely to reject first?

Scale:
1 = "Much more likely to reject A", 2 = "More likely to reject A", 3 = "Slightly more likely to reject A," 4 = "About the same", 5 = "Slightly more likely to reject B", 6 = "More likely to reject B", 7 = "Much more likely to reject B"
(Note: Given study 4b is in a negative domain we focus on the Confidence measure in the manuscript.)

Helpfulness:

- How helpful were the posts of BBQ A relative to BBQ B? (No photo condition)
- How helpful were the posts with photos in making your decision between BBQ A and B? (Photo condition)

Scale:
1 = "Much less helpful", 2 = "Less helpful", 3 = "Moderately less helpful," 4 = "About the same", 5 = "Moderately more helpful", 6 = "More helpful", 7 = "Much more helpful"

(Note: Results of the *Helpfulness* measure were significant and in the direction expected, but these results were not included in the manuscript for the purpose of brevity.)

Confidence

- One BBQ pit is more popular than the other: Which one are you less confident of in its popularity?
- One BBQ pit is recommended by more customers than the other: which one are you less confident in its recommendations?
- For one BBQ pit, customers are more satisfied with than the other: which one are you less confident in its satisfaction rate?

Scale:
1 = "A lot less confident in A", 2 = "Less confident in A", 3 = "Slightly less confident in A," 4 = "About the same", 5 = "Slightly less confident in B", 6 = "Less confident in B", 7 = "A lot less confident in B"

Appendix I: Study 5: Serial Mediation Confidence in Massage Gun Reviews

Below, are several features of the Massage gun. You can read them quickly:

- Effectively Relieve Muscle Soreness: massage gun uses percussion therapy, which efficiently helps relax your stiff muscles, pain, lactic acid build-up, improves blood flow, and accelerates body recovery. Perfect for athletes, the office sedentary, the elders, fitness instructors and hiking.

- Multiple Speeds & Massage Heads: professional deep tissue massage gun provides multiple vibration mode. Suitable for the neck, shoulder, arms, back, waist, leg, feet, and other muscle groups.
- High Quality & Ultra-Quiet Massager: The build of the muscle gun is solid, equipped with a high-torque brushless motor and noise reduction technology.
- Long Working Time: Rechargeable battery, 8 hours battery life. The charging time is 4 hours. It is designed with a 15-minute auto-off setting to prevent overheating to extend the life of the battery.
- Easy to Hold & Use: Easy to use, non-slip and anti-drop handle, lightweight.

Photos and description provided by the brand:

WESKEAN	TYIAUS
	

Photos included with the reviews:

Measures

Relative choice likelihood:

- Suppose that you only had these posts to make your decision. How likely would you be to pick TYIAUS massage gun vs WESKEAN massage gun?

Scale:
1 = “Much more likely WESKEAN”, 2 = “More likely WESKEAN”, 3 = “Slightly more likely WESKEAN,” 4 = “About the same”, 5 = “Slightly more likely TYIAUS”, 6 = “More likely TYIAUS”, 7 = “Much more likely TYIAUS”.

Helpfulness:

- Which post did you find more helpful?

Scale:
1 = “Much more helpful Adam”, 2 = “More helpful Adam”, 3 = “Moderately more helpful Adam,” 4 = “About the same”, 5 = “Moderately more helpful Bob”, 6 = “More helpful Bob”, 7 = “Much more helpful Bob”

Confidence in writer:

- Which reviewer has signalled more confidence with their post?
- Which reviewer has displayed more belief in their post?
- Which reviewer has shown that they stand by their post more?

Scale:
100 point slide scale. Labelled: 1 = “A lot more Adam”, 2 = “More Adam”, 3 = “About the same”, 4= “More Bob”, 5= “A lot more Bob”.

Confidence of reader:

- For which Massage gun are you more confident that it is popular with customers?
- For which Massage gun are you more confident that customers would recommend it to others?
- For which Massage gun are you more confident that customers are satisfied with it?

Scale:
1 = “Much more confident in WESKEAN”, 2 = “More confident in WESKEAN”, 3 = “Slightly more confident in WESKEAN,” 4 = “About the same”, 5 = “Slightly more confident in TYIAUS”, 6 = “More confident in TYIAUS”, 7 = “Much more confident in TYIAUS”.