THE ROLE OF ONLINE REVIEWS IN CONSUMER DECISION-MAKING

A Dissertation
Presented to
The Academic Faculty

by

Zhanfei Lei

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Scheller College of Business

Georgia Institute of Technology
August 2019

THE ROLE OF ONLINE REVIEWS IN
CONSUMER DECISION-MAKING

Approved by:

Dr. Han Zhang, Co-Advisor
Scheller College of Business
Georgia Institute of Technology

Dr. Eric Overby
Scheller College of Business
Georgia Institute of Technology

Dr. Dezhi Yin, Co-Advisor
Muma College of Business
University of South Florida

Dr. Sridhar Narasimhan
Scheller College of Business
Georgia Institute of Technology

Dr. Sabyasachi Mitra
Scheller College of Business
Georgia Institute of Technology

Date Approved: July 10, 2019

To my parents

&

my husband Yonghao

ACKNOWLEDGEMENTS

I am grateful to my advisors, Dr. Han Zhang and Dr. Dezhi Yin, for their guidance and mentorship. I want to forever thank Han for guiding me into the academia world, showing me how to become a rigorous scholar, and providing me with tremendous support and patience. I am also sincerely grateful to Dezhi for generously sharing his knowledge and valuable experience, advising me to read and think, and always being a source of inspiration and encouragement. Thanks Han and Dezhi for everything. I am so fortunate to work with them and have them as my advisors.

I would like to thank the rest of my committee members, Drs. Saby Mitra, Eric Overby, and Sridhar Narasimhan, who provide constructive feedback and insights for my thesis. I also owe many thanks to Drs. Marius Florin Niculescu, Lizhen Xu, and Michael Smith for recruiting their students to participate in my experiments. I thank Drs. D.J. Wu, Jeffery Hu, and Mingfeng Lin for their excellent comments and suggestions on my work over years. Many special thanks to my student cohort at Scheller for their support and encouragement.

Finally, I wish to express my deepest appreciation to my family and my friends. I want to thank my parents, for their unconditional love and support, and for having faith in me. They are also my mentors, who are always standing behind me. I am truly indebted to my husband and my best friend, Yonghao Yu, for his endless love and understanding. I really thank him for always being supportive, without whom I could not have completed this journey. Last, many thanks to Yanran Liu, Xinying Liu, Yukun Yang, Ming Jiang, and those I do not mention. I feel so lucky to have them as my friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

As a prominent form of user-generated content, online reviews have become increasingly indispensable for consumers to make purchase decisions. Thus, it is critical to understand what sets helpful reviews apart from unhelpful ones, which kinds of reviews consumers prefer to read, and how online reviews shape consumers' purchase decisions. Prior research has examined diverse determinants of review helpfulness and the effect of summary rating statistics on product sales. However, few studies have examined the impact of reviewers' writing styles on review helpfulness, explored the critical role of consumers' initial beliefs before they read and evaluate reviews, or investigated the likelihood of individual reviews to sway consumers' purchase decisions. Addressing these important gaps and scrutinizing commonly accepted assumptions, my dissertation aims to explore how, why, and when various aspects of online reviews influence consumers' judgment of review helpfulness, preference in information seeking, and purchase decisions.

In the first essay, I explore how and when reviewers' use of personal pronouns, a type of "invisible" words that do not carry substantive meanings, can influence reader evaluation of reviews. Drawing on the empathy and persuasion literature, I develop a theoretical framework suggesting that personal pronouns may influence consumers' perceived review helpfulness through opposing processes, and that the overall effect of personal pronouns is contingent on the review's two-sidedness. Results of five studies (one archival analysis and four controlled experiments) provide consistent support for the proposed hypotheses. The findings deepen understanding of the role of "invisible" words in online reviews, reveal an important boundary condition for the impact of personal

pronouns, and provide practical implications for product/service providers, review platforms and reviewers.

In the second essay, I investigate consumers' preference for confirmatory or disconfirmatory reviews during the information seeking and evaluation stages of their decision-making process. Drawing on the motivated reasoning literature, I propose that consumers with positive initial beliefs about a product tend to seek out disconfirmatory reviews (termed disconfirmation bias), while consumers with negative initial beliefs tend to search for confirmatory reviews (termed confirmation bias). However, consumers would evaluate confirmatory reviews more favorably regardless of the valence of their initial beliefs. Results of three controlled experiments provide consistent support for these hypotheses. The findings reveal the occurrences of differential biases in different stages of decision-making, deepen understanding of consumers' initial beliefs, and provide important implications for product manufacturers and review platforms.

In the third essay, I compare the relative impact of a product's overall average rating and its prominently displayed individual reviews. Drawing on the heuristics-and-biases literature, I derive two competing hypotheses that contrast the former's anchoring effect and the latter's swaying effect. In addition, extending the concept of belief confidence from the metacognition literature, I propose the dispersion of the product's ratings as a boundary condition for the anchoring and swaying effects. Through two carefully designed experiments and an archival analysis of a panel dataset collected from Apple's App Store, I find evidence that consumers' intention to purchase the product and product sales are influenced primarily by the most accessible reviews (i.e., swaying effect), and that rating dispersion can moderate both swaying and anchoring effects. The findings challenge the

common belief of researchers and practitioners that the most critical determinant of product

sales is the average product rating, and offer important theoretical and practical

implications.

CHAPTER 1.   INTRODUCTION

Online reviews containing judgement and opinions from previous consumers play an increasingly important role in prospective consumers' purchase decisions. More and more consumers choose to read online reviews to get sufficient knowledge of the product before making a purchase decision. However, the exploding amount of online review information makes it difficult for consumers to identify useful and necessary information efficiently, resulting in information overload (Jones et al. 2004). Therefore, to account for information overload and reduce consumers' effort and time in decision-making, review platforms incorporate various strategies dealing with online reviews, such as providing voting systems to evaluate the helpfulness of reviews, allowing consumers to customize the sort order of reviews based on their preference, and displaying text reviews as well as summary rating statistics for consumers.

Review platforms' diverse strategies have attracted great attention from researchers. For instance, prior studies have explored antecedents of review helpfulness, including numerical product ratings, product and reviewer characteristics, review timing, and characteristics of review text (e.g., Forman et al. 2008; Korfiatis et al. 2008; Mudambi and Schuff 2010; Chen and Lurie 2013). In particular, the identification of review text characteristics, such as information amount (Kim et al. 2006), readability of the text (Korfiatis et al. 2008), and emotional expressions (Yin et al. 2014; Yin et al. 2017), was primarily based on content words that have real meanings. Few studies explored the influence of function words such as personal pronouns on review helpfulness. In addition, prior research generally assumed that different consumers would assess the same review

similarly, while resent studies have demonstrated that consumers' initial beliefs about a product play a role in their helpfulness evaluation of product reviews (Cheung et al. 2009; Qiu et al. 2012; Yin et al. 2016). In addition to affect consumers' evaluation of online reviews, their initial beliefs could also be influential in the way they search for online reviews, which is worthy of further investigation. Moreover, although both summary rating statistics such as the overall average rating and individual reviews can affect consumers' purchase decision and it is generally assumed that the former matters more than the latter for producing product sales (e.g., Shoham et al. 2017; Babić Rosario et al. 2016; Floyd et al. 2014), little is known about the direct comparison between the impact of aggregated rating statistics with that of individual reviews on product sales.

Addressing these important gaps and revisiting commonly accepted assumptions, this dissertation aims to examine how, why, and when different aspects of online reviews impact consumers' helpfulness evaluation of online reviews, preference in information seeking, and intention to purchase the product.

1.1 Essay 1

In the first essay, I explore how and when reviewers' use of first-person vs. second-person pronouns in expressing their opinions impacts prospective consumers' perceptions of review helpfulness. Prior studies on the online review literature have revealed a number of factors that influence review helpfulness, including product ratings, product and reviewer characteristics, and characteristics of review text (e.g., Forman et al. 2008; Huang et al. 2015; Mudambi and Schuff 2010; Korfiatis et al. 2008; Yin et al. 2014). In particular, the majority of studies on examining the role of review text characteristics was based on

2

content words (e.g., nouns, verbs, adjectives) that have real meanings. However, few studies explored the impact of "invisible" and "forgettable" function words (e.g., pronouns, conjunctions, and prepositions) that do not convey substantive meanings but reveal how people use words to convey the message (Campbell and Pennebaker 2003). In addition, as a particular category of function words, personal pronouns (e.g., "I", "you") have been demonstrated to reflect one's focus of attention, such that first-person pronouns (e.g., "I", "me", "my") indicate one's attention toward oneself, and second-person pronouns (e.g., "you", "your") indicate one's attention toward others (Ickes et al. 1986; Pennebaker et al. 2003).

Since a focus on others (vs. self) is influential in offline interpersonal settings even between strangers (Fraley and Aron 2004; Galinsky and Moskowitz 2000), reviewers' use of personal pronouns that reflects their attentional focus could play a nontrivial role in review readers' helpfulness evaluation of reviews. Drawing on the empathy and persuasion literature (Campbell and Kirmani 2000; Hodges et al. 2011), I develop a theoretical framework suggesting that reviewers' use of personal pronouns may play a role in consumers' perception of review helpfulness via both positive and negative processes, such that consumers are likely to perceive a reviewer using more second-person (vs. first-person) pronouns to be more empathic concern for them (Goldstein et al. 2014), and they are also likely to infer that the reviewer intentionally uses tactics to persuade them (Kirmani and Campbell 2004). In addition, I propose that review two-sidedness can be a boundary condition for the overall effect of personal pronouns, such that the positive effect of second-person (vs. first-person) pronouns is attenuated for one-sided reviews.

To test proposed hypotheses, I conduct an archival analysis and four controlled experiments. In Study 1, utilizing an archival dataset collected from Apple's App Store, I measure the ratio of second-person pronouns in online reviews and explore the overall effect of personal pronouns on review helpfulness. In Study 2, to account for alternative explanations that cannot be ruled out in Study 1 and examine the moderating role of review two-sidedness, I conduct an experiment in which I directly manipulate personal pronouns (first-person pronouns only vs. second-person pronouns only) and review two-sidedness (one- vs. two-sided). In Study 3, I investigate the moderating role of review two-sidedness by varying its level (low vs. high two-sidedness) and examine the mechanisms underlying the main effect. In Study 4A and Study 4B, I explore the role of personal pronouns for both positive and negative one-sided reviews, and examine the underlying mechanisms. The five studies provide converging evidence for the nontrivial role of reviewers' use of personal pronouns in review readers' helpfulness perception of reviews, and reveal that the positive effect of second-person (vs. first-person) pronouns is greater for two-sided reviews compared with one-sided reviews.

These findings contribute to the online review and linguistic literature. First, this essay goes beyond content words with real meanings, reveals that reviewers' unconscious use of personal pronouns can play a role in review readers' perception of reviews, and deepens the understanding of the influence of "invisible" function words on online reviews. Second, the findings add to the linguistic literature by extending the study of personal pronouns from offline settings typically involving two individuals to an online environment involving hundreds of thousands of prospective consumers who are total strangers during a persuasion process. This essay also provides implications for review

platforms and reviewers, such that switching from "I" to "you" could be an efficient strategy of writing reviews to boost review helpfulness, especially when reviewers have mixed opinions.

1.2 Essay 2

In the second essay, I explore consumers' preference for confirmatory or disconfirmatory reviews when they seek and evaluate review information about a product. While prior research on online reviews generally assumed that different consumers assess the same review in a similar way, recent studies have started to question this assumption by identifying the role of consumers' initial beliefs in their evaluation of reviews (Cheung et al. 2009; Qiu et al. 2012; Yin et al. 2016). In particular, Yin et al. (2016) found that consumers evaluate confirmatory reviews consistent with their initial beliefs as more helpful than disconfirmatory reviews, resulting in a confirmation bias. Because consumers generally have positive initial beliefs driven by the positive average rating of most products, overall they should demonstrate a positivity bias by evaluating positive (confirmatory) reviews as more helpful than negative (disconfirmatory) reviews. However, existing literature has repeatedly revealed a negativity bias in the influence of online reviews on product sales – negative reviews hurt product sales to a greater extent than positive reviews help product sales (e.g., Basuroy et al. 2003; Cao et al. 2011; Chevalier and Mayzlin 2006). It is quite interesting to uncover the probable reasons underlying this dilemma: negative reviews are generally evaluated as less helpful than positive reviews while negative reviews have greater influence on product sales than positive reviews. One possible reason is that the way consumers look for information also matters, because they need to be selective in looking for numerous online review information to read (Mathieson and Wall

1982; Woodside and MacDonald 1994). Therefore, it is crucial to examine consumers' preference for confirmatory or disconfirmatory information that is contingent on their initial beliefs when they seek out online reviews to read – the earlier stage of consumers' decision-making process.

Drawing on the motivated reasoning literature, I propose two hypotheses of consumers' preference for confirmatory vs. disconfirmation reviews during differential stages of their decision-making process. I propose that in the information seeking stage, consumers with positive initial beliefs prefer to look for disconfirmatory reviews, while consumers with negative initial beliefs prefer to seek out confirmatory reviews. In the information evaluation stage, consumers would evaluate confirmatory reviews more favorably regardless of the valence of their initial beliefs. To test these hypotheses, I conduct three controlled experiments. In Study 1, I design an experiment in which participants form positive initial beliefs about a product before seeking for and evaluating online reviews. In Study 2, based on a more realistic scenario, I replicate findings of Study 1. In Study 3, I manipulate the valence of participants' initial beliefs and examine the whole story. The results demonstrate a general negativity bias in the information seeking stage – consumers with positive initial beliefs tend to seek disconfirmatory (negative) reviews while consumers with negative initial beliefs tend to seek confirmatory (negative) reviews; in the information evaluation stage, consumers generally prefer confirmatory reviews compared with disconfirmatory reviews, leading to a confirmation bias.

This essay has several theoretical and practical implications. First, this research is among the first attempts to examine consumers' information seeking behavior during an earlier stage of consumers' decision-making process. Second, by scrutinizing the role of

consumers' initial belfies during differential stages of the decision-making process, this essay provides a possible reason for the negativity bias demonstrated in the existing literature. The results reveal that negative reviews may not be always more helpful than positive reviews, and that consumers' greater exposure of negative reviews may overweigh the greater helpfulness of positive reviews, leading to a negativity bias regarding the impact of online reviews on product sales. In addition, a deeper understanding of how consumers look for information can provide implications for practitioners to sort and deal with online reviews considering their potential exposure to consumers.

1.3 Essay 3

In the third essay, I examine how and when online reviews can sway consumers' attitude toward a product. To help consumers efficiently gauge product quality, many online retailers and review platforms display summary ratings such as the overall average rating in a more prominent place than individual reviews. In particular, the overall average rating incorporating comprehensive product information is the most important signal of product quality (De Langhe et al. 2015), and helps consumers to form a first impression of a product (Yin et al. 2016). For researchers, it is also an implicit assumption that aggregated rating statistics are more influential than individual reviews in predicting product sales, although they both have been demonstrated to influence consumers' purchase decisions (e.g., Babić Rosario et al. 2016; Floyd et al. 2014; Shoham et al. 2017). However, no research has directly compared the impact of aggregated rating statistics with the most accessible individual reviews on product sales. If consumers rely more on individual reviews during the decision-making process and individual reviews can easily sway

product sales, practitioners' emphasis on the aggregated rating statistics such as the overall average rating might be misguided.

In this essay, I compare the relative influence of a product's overall average rating and its most accessible individual reviews on consumers' purchase decisions. Building on the heuristics-and-biases literature (Tversky and Kahneman 1974), I propose two competing hypotheses. On one hand, the overall average rating may be more influential than individual reviews (termed as the anchoring effect) because it serves as a salient anchor and leads consumers' final judgment to be closer to the anchor. On the other hand, individual reviews may play a greater role in consumers' purchase decisions and product sales (termed as the swaying effect) because they are more cognitively accessible and perceived as more important for consumers. In addition, drawing on the belief confidence literature (Smith and Swinyard 1988), I posit the rating distribution can be a boundary condition for the anchoring and swaying effects.

To test these hypotheses, I conduct two designed experiments and an archival analysis. In Study 1, I conduct an experiment to examine the existence of the anchoring versus swaying effects by directly manipulating the overall average rating of two products and their most recent reviews. In Study 2, I design a follow-up experiment in which I extend the manipulation of rating distributions and explore the boundary condition for the anchoring and swaying effects. In Study 3, utilizing a panel dataset collected daily from Apple's App Store over two months, I replicate the findings of Study 1 and Study 2. The results of three studies provide converging evidence for a swaying effect – individual reviews are more influential than the overall average rating in consumers' purchase decisions, and for the moderating role of the rating distribution.

This essay has a number of contributions. First, the findings challenge the commonly accepted assumption that the overall average rating matter more than induvial reviews for predicting product sales, and reveal a swaying effect – individual reviews that are the most accessible for consumers can easily sway their purchase intentions. Second, extending insights of heuristics-and-biases literature into consumer behavior research, this essay indicates that the availability heuristic is more applicable than the anchoring heuristic in the context of online reviews. In addition, the findings suggest that product manufactures and online retailers can better influence consumers by paying more attention to the first reviews that consumers are likely to see and taking a more balanced view between the summary ratings and individual reviews.

CHAPTER 2.   WHEN "I" BECOMES "YOU": THE ROLE OF

PERSONAL PRONOUNS IN ONLINE REVIEWS

2.1   Introduction

As a prominent form of user-generated content, online reviews have become increasingly indispensable for consumers to make purchase decisions. However, the exploding number of online reviews can appear overwhelming for consumers and cause information overload (Jones et al. 2004). Thus, most online review platforms identify and promote helpful reviews by asking review readers to vote on the helpfulness of reviews and then displaying helpful ones more prominently. A deeper understanding of what factors contribute to helpful reviews has clear benefits to product/service providers, review platforms, and reviewers (Mudambi and Schuff 2010).

Following prior literature, we define review helpfulness as the extent to which an online review is perceived by consumers to facilitate their purchase decision process (Yin et al. 2014). Review helpfulness reflects perceived value or diagnosticity of review information, as the review can provide diagnostic value for consumers' judgment and decision-making process (Mudambi and Schuff 2010). Prior studies have examined a variety of factors that influence review helpfulness, including ratings, product type, reviewer characteristics, and consumers' initial beliefs (e.g., Forman et al. 2008; Huang et al. 2015; Mudambi and Schuff 2010; Yin et al. 2016). The characteristics of review text have also been demonstrated as influential determinants, such as information amount (Kim

et al. 2006), readability of the text (Korfiatis et al. 2008; Krishnamoorthy 2015), and emotional expressions (Yin et al. 2014; Yin et al. 2017).

The identification of these text characteristics was based primarily on content words that have real meanings (e.g., nouns, verbs, adjectives). However, few studies examined the role of "invisible" function words. Function words are words that serve grammatical purposes (e.g., pronouns, conjunctions, and prepositions) and connect content words. These "invisible" words do not convey substantive meanings like content words, and they account for only a very small percentage of all the words (Corver and van Riemsdijk 2001). However, function words reflect people's personality and psychological states and have nontrivial implications for individuals (Pennebaker 2011). For example, the use of a particular category of function words, personal pronouns (e.g., "I", "you"), has been revealed to influence health outcomes, social relationships, as well as stereotypic biases and prejudice (e.g., Arriaga and Rusbult 1998; Campbell and Pennebaker 2003; Galinsky and Moskowitz 2000).

In this work, we explore the implications of personal pronouns in online reviews. Personal pronouns appear frequently in text reviews, but they are largely invisible and forgettable for review readers. However, the use of personal pronouns indicates an individual's focus of attention, such that first-person pronouns (e.g., "I") reflect one's attention toward oneself, and second-person pronouns (e.g., "you") reflect attention toward others (Ickes et al. 1986; Pennebaker et al. 2003). Because a focus on others (vs. self) plays an important role in offline interpersonal settings even between strangers (Fraley and Aron 2004; Galinsky and Moskowitz 2000), the use of second-person (vs. first-person) pronouns that indicate reviewers' attentional focus may have a similarly nontrivial impact on review

readers' perception of reviews. As a motivating example, consider the following two hypothetical reviews of a restaurant:

Reviewer A: *"As soon as I walk in, I can't help but feel a positive vibe in the air. The owner is energetic, friendly, humble, and treats me like family. The main entrees here are probably the best dish I've ever had in my life."*

Reviewer B: *"As soon as you walk in, you can't help but feel a positive vibe in the air. The owner is energetic, friendly, humble, and treats you like family. The main entrees here are probably the best dish you've ever had in your life."*

Reviewer A uses first-person pronouns to describe the experience with a focus on himself/herself, whereas Reviewer B uses second-person pronouns to describe the same experience with a focus on others (prospective consumers). Which of the two reviews will be considered more helpful by prospective consumers? More broadly, does reviewers' use of second-person (vs. first-person) pronouns influence consumers' perception of review helpfulness, and under what conditions?

To answer these questions, we draw on the empathy and persuasion literature (Campbell and Kirmani 2000; Hodges et al. 2011), and theorize that reviewers' use of second-person (vs. first-person) pronouns can influence readers' perception of review helpfulness via both positive and negative processes: review readers may perceive a reviewer using more second-person (relative to first-person) pronouns to be more empathic for and concerned about them (see Goldstein et al. 2014), and they may also infer the reviewer to have a hidden intent to persuade them (see Kirmani and Campbell 2004). We further propose that the role of reviewers' use of personal pronouns is more nuanced than

expected, and identify two-sidedness as a boundary condition for the overall effect of personal pronouns (e.g., the opposing processes may cancel out each other in one-sided reviews).

To test this theoretical framework, we conducted five studies with distinct methodologies, including one archival analysis and four controlled experiments. Our research makes two primary contributions. First, a growing literature has examined how characteristics of review text influence consumer perception of reviews (e.g., Jensen et al. 2013; Yin et al. 2014). While the existing studies have been concerned with various aspects or types of content words such as expressed emotions, much less is known about the role of function words such as personal pronouns, which are largely "invisible" but at the same time more fundamental elements of verbal language (Pennebaker 2011). We add to this literature by showing that reviewers' unconscious use of "invisible" personal pronouns can influence reader perceptions of the reviews. Specifically, we find evidence that the use of second-person (vs. first-person) pronouns can lead readers to perceive both empathic concern and persuasion motives from the reviewers, which have opposite implications for the consumers' evaluation of review helpfulness. Thus, this paper takes an initial step to go beyond content words conveying substantive meanings and deepens our understanding of the role of "invisible" words in online reviews. Second, our findings add to the linguistic literature by extending the study of personal pronouns from offline settings typically involving two individuals to an online environment involving hundreds of thousands of prospective consumers who are total strangers during a persuasion process. In particular, we examine a boundary condition in this unique persuasion context that can change the relative power of the two opposing processes underlying the overall effect of personal

pronouns. While the use of second-person (vs. first-person) pronouns has been found to facilitate various positive outcomes in offline contexts (e.g., Coke et al. 1978; Galinsky and Moskowitz 2000; Hodges et al. 2011), we showed across multiple studies that personal pronouns are not likely to influence review helpfulness when the reviews are one-sided with only praises or critiques. This finding reveals that the impact of personal pronouns is contingent on the two-sided nature of reviews and deepens our understanding of boundary conditions for this impact. Our findings also provide clear practical implications for product/service providers, review platforms, and reviewers.

2.2 Theoretical Framework and Hypotheses

2.2.1 Function Words and Personal Pronouns

Words can be categorized into content words and function words. Content words, such as nouns, verbs, and adjectives, have real meanings and reflect what people are talking about. Function words, such as pronouns, prepositions, and conjunctions, do not convey substantive meanings but reveal how people use words to convey the message (Campbell and Pennebaker 2003). Although these often overlooked function words account for a trivial percentage (less than 0.04%) of the total vocabulary (Chung and Pennebaker 2007) and they seem "invisible" and "forgettable", the use of these function words can reflect people's personality and psychological states (Pennebaker 2011). Thus, the relationship between function words and human psyche should not be ignored.

As a commonly encountered category of function words, personal pronouns (e.g., "I", "you") have been revealed to serve important psychological functions for individuals. For instance, the use of first-person singular pronouns in people's political speeches and

medical interviews is closely associated with their tendency to be depressed (Weintraub 1989). Pronouns were found to be a better signal of depression compared with negative emotion words, and depressed students used more first-person singular pronouns in their thoughts about coming to college than formerly or never depressed students (Rude et al. 2004). In addition, leaders' (such as a mayor's) use of first-person plural pronouns in speeches can reveal their close emotional ties to and strong social bounds with others (Pennebaker and Lay 2002).

In verbal communication, the use of personal pronouns can also indicate one's focus of attention. In contrast to self-focused attention, an individual's other-focused attention refers to one's attention focused on thoughts and feelings of others rather than on one's own (Ingram 1990; Mor and Winquist 2002). Previous research showed that greater use of first-person pronouns (e.g., "I," "me," "my") reflects the self-focus of an individual (Pennebaker et al. 2003), and that greater use of second-person pronouns (e.g., "you," "your") reflects the other-focus (Ickes et al. 1986; Simmons et al. 2005). For example, Simmons et al. (2005) showed that individuals who use more second-person and less first-person pronouns are more likely to consider a situation from another's viewpoint, thus being more other-focused.

A focus of attention on others plays a key role in social interactions and relations. The ability and propensity to consider other people's thoughts and feelings has long been recognized as an invaluable tool for proper social functioning (Galinsky and Moskowitz 2000; Higgins 1981; Todd et al. 2011). Prior literature has revealed that other-focus serves as a primary developmental breakthrough in cognitive functioning (Piaget 1932), plays an important role in moral reasoning (Kohlberg 1976), correlates with social competence and

social esteem (Davis 1983), and reduces stereotypic biases and prejudice (Galinsky and Moskowitz 2000). Thinking about others has also been shown to increase an individual's empathy, willingness to help others (Coke et al. 1978; Toi and Batson 1982), and other gestures of altruism (Batson 1991; Batson 1998). People in close relationships who make more effort in focusing their attention on their romantic partners also develop more positive attitude toward their partner and the relationship (Arriaga and Rusbult 1998; Long and Andrews 1990). Overall, these findings support a predominant view that shifting one's attention from oneself to others is beneficial and leads to positive outcomes for oneself (Hodges et al. 2011).

Because of the close association of attentional focus and personal pronouns, the greater use of second-person (relative to first-person) pronouns should also lead to positive outcomes for individuals. Despite what we learned about personal pronouns from the literature on linguistics and attentional focus, the unique nature of consumer reviews warrants further investigation of the role of personal pronouns in the online context. First, second-person pronouns and other-focus have been found to play a significant role in people's prosocial behaviors (e.g., Batson et al. 1997; Galinsky and Moskowitz 2000) and in competitive environments such as negotiations (e.g., Galinsky et al. 2008). However, much less is known about their role in persuasion contexts despite the importance and prevalence of the use of personal pronouns in persuasion (Clark and Delia 1976). Because the primary purposes of online reviews are to inform and persuade future consumers (Sparks et al. 2013), more research is needed to understand the impact of personal pronouns in this persuasion process. Second, the very few studies looking into personal pronouns in online reviews provided suggestive evidence that the role of personal pronouns might be

more nuanced in this unique context. For instance, Schindler and Bickart (2012) found that online reviews contain significantly more first-person pronouns than second-person pronouns, but the use of personal pronouns did not impact review helpfulness. Another exploratory analysis of Amazon reviews showed that helpful reviews contain less first-person singular pronouns than unhelpful reviews, but the percentage of second-person pronouns did not differ significantly (Liang et al. 2014). These descriptive findings suggest the need to systematically examine whether, how, and when personal pronouns can influence perceived review helpfulness. In the following, we develop a theoretical framework based on the empathy and persuasion literature.

2.2.2   *Perceived Empathic Concern*

Extending the prior literature studying the association of other-focus and empathy (Coke et al. 1978; Toi and Batson 1982), we propose that a reviewer's use of second-person (vs. first-person) pronouns that reflects one's attentional focus on others can positively influence readers' perception of review helpfulness through their perception of the reviewer's empathic concern (an important component of empathy; see Hodges et al. 2011). [1]

Empathic concern refers to one's concerns or compassion for others (Batson 1987).[2]   A focus of attention on others has been well established as a reliable means of

---

[1] We acknowledge that the positive effect of personal pronouns on review helpfulness might also be explained by factors other than perceived empathic concern. For example, reviewers who use more second-person pronouns may be perceived to be altruistic, willing to help others and share their expertise with others. We investigate these alternative explanations further in a supplementary study reported in Appendix B.

[2] Sympathy and empathic concern are both emotional responses to others' feelings, but they are distinct psychological processes (Wispé 1986). Sympathy refers to one's heightened awareness of another's distress and suffering, whereas empathic concern captures one's absorption in the feelings of another (Escalas and Stern 2003). It is the latter that is more relevant in our context.

activating one's empathic concern for others (Batson 2009; Hoffman 2001). Because it is intuitive and widely known that other-focus allows one to better appreciate the situations that the targets on the receiving end are encountering, the targets are likely to make this connection and perceive the other-focused individual as feeling more empathic concern for them (e.g., Batson et al. 1996; Hodges et al. 2010). Goldstein et al.'s (2014) study also provided direct evidence for this association between perceived other-focus and perceived empathic concern in explaining people's prosocial behavior. Applied to our setting, when consumers read a review containing more second-person (vs. first-person) pronouns, they become targets of the review writer's attentional focus and are likely to perceive the reviewer as more empathic for them and more concerned about them.

Review readers' perception of greater empathic concern from a reviewer should in turn lead them to perceive the review as more helpful. When people are aware of another person's concern about their well-being, they tend to trust the person because of his/her kindness (Johnson et al. 1996; Mayer et al. 1995) and develop positive feelings toward the person (e.g., Newcomb 1956). Such a favorable impression of another person and the associated positive feelings can spill over to other aspects of the communication process, leading people to evaluate communicated information more favorably (Pornpitakpan 2004). In our setting, review readers who perceive greater empathic concern from a reviewer should trust and like the reviewer to a greater extent. Combined with recent evidence that reviews from more credible sources are perceived more helpful (Baek et al. 2012; Cheung et al. 2012), it is reasonable to expect review readers to associate greater empathic concern with more helpful reviews. Taken together, we propose the following hypothesis.

Hypothesis 1: A reviewer's use of second-person (vs. first-person) pronouns in a review is positively associated with the helpfulness of the review perceived by readers.

While the positive association between the use of second-person (relative to first-person) pronouns that reflect other-focused attention and perceived review helpfulness makes intuitive sense, it is unclear whether this positive association always holds, and whether second-person pronouns can trigger negative reactions on the part of review readers under certain conditions. Most studies on attentional focus have emphasized a positive effect of other-focus in diverse contexts (Hodges et al. 2011). However, recent research has started to question this assumption and speculated that other-focus may also backfire and negatively affect other-focused individuals under certain situations (Sassenrath et al. 2016; Vorauer 2013). Extending the persuasion knowledge model (Friestad and Wright 1994) in the next section, we argue that review readers may perceive a reviewer who uses more second-person pronouns as having a higher intent to persuade and subsequently evaluate the review less favorably, and that this is more likely to occur when the review is one-sided than two-sided.

2.2.3   Perceived Persuasion Motives and Review Two-Sidedness

Review readers can also interpret a reviewer's use of second-person (vs. first-person) pronouns in ways that reduce their perception of review helpfulness, and one such interpretation is an inference of the reviewer's persuasion motives. The writing and reading of online reviews are essentially a persuasion process (Sparks et al. 2013), in which the reviewer (the "source" or persuasion agent) writes a review (the "message") to persuade prospective readers (the "recipient"). Specifically, we define the inference of persuasion

motives as review readers' inference that the reviewer has a hidden intent to persuade readers (see Campbell and Kirmani 2000). According to the persuasion knowledge model, when people are exposed to a persuasion agent's persuasion tactics, this exposure can activate people's belief that the agent has persuasion motives—using the tactics to persuade (Friestad and Wright 1994; Kirmani and Campbell 2009). Such a belief is likely to be triggered if the agent is known to benefit from the tactics (Campbell and Kirmani 2000; Laran et al. 2011). The use of second-personal pronouns in online reviews indicates reviewers' other-focused attention, which has been commonly associated with positive outcomes for other-focused individuals (Hodges et al. 2011). As a result, when consumers read a review using more second-person (and less first-person) pronouns, they are more likely to infer that the reviewer intentionally uses this tactic to persuade them.

In addition, we propose that the likelihood of this negative inference depends critically on whether the review is one-sided or two-sided. Review two-sidedness refers to the extent to which a review consists of information about both positive and negative attributes of a product (Jensen et al. 2013). Previous studies showed that two-sided reviews are viewed as more credible and trustworthy than one-sided reviews (e.g., Jensen et al. 2013). Reviewers crafting fake reviews are most likely to compose an extreme review rather than a two-sided review in order to benefit the company/business or competitor (Luca and Zervas 2016). In contrast, reviewers commenting on both sides are more likely to be objective and tell the truth about the product than using manipulative tactics to persuade review readers, because two-sided information would be treated as more accurate and representative of the truth of the product (Crowley and Hoyer 1994; Jensen et al. 2013).

Hence, negative inferences of reviewers based on the use of their personal pronouns are more likely to occur for one-sided reviews than two-sided reviews.

The perception of greater persuasion motives from a one-sided reviewer can in turn reduce review readers' perception of review helpfulness. Based on reactance theory, inferences of another's persuasion motives result in a feeling of pressure and a potential threat for freedom, which can in turn lead to more resistance to being persuaded (Brehm and Brehm 2013; Clee and Wicklund 1980). Accumulating evidence from diverse contexts also suggests that the perception of an influence agent's use of manipulative tactics undermines message persuasiveness (see Sagarin et al. 2002). In our context, review readers typically have the freedom to hold whatever opinions they have about a product or read whatever reviews they find helpful. Thus, the readers' perception of a reviewer intending to influence their choice can activate their reactance, leading them to lower their perception of review helpfulness.

To summarize, the negative process underlying the impact of reviewers' use of second-person (vs. first-person) pronouns through inferences of their persuasion motives is stronger for one-sided than two-sided reviews, but the positive process through perceived empathic concern should remain the same given the close association of other-focus with empathy. Therefore, we expect the positive effect of second-person (vs. first-person) pronouns to be attenuated for one-sided reviews and propose the following hypothesis. Our research model is shown in Figure 2-1.

Hypothesis 2: The positive effect of a reviewer's use of second-person (vs. first-person) pronouns on perceived review helpfulness is greater for two-sided reviews compared with one-sided reviews.
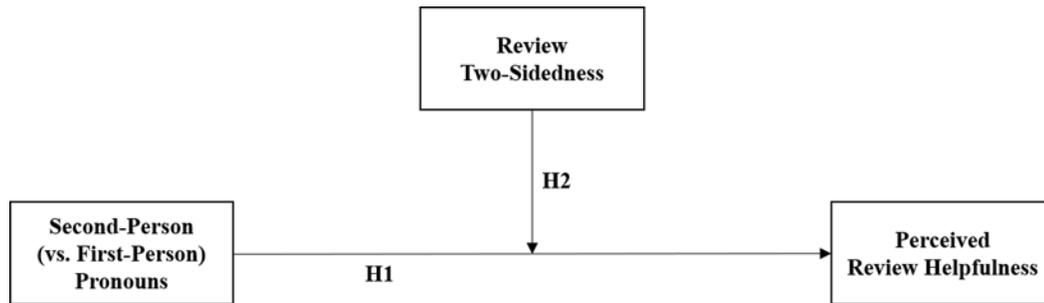
Figure 2-1 Research Model.

To test these hypotheses, we conducted five studies with distinct methodologies. The first two studies examined the direct impact of personal pronouns on review helpfulness and the moderating role of review two-sidedness, while the following three studies probed into the probable mechanisms underlying the effect of personal pronouns. In Study 1, we utilized an archival dataset collected from Apple's App Store, measured the ratio of second-person pronouns in text reviews, and tested its main effect on review helpfulness (H1). In Study 2, we conducted an experiment, in which we directly manipulated personal pronouns (first-person pronouns only vs. second-person pronouns only) and review two-sidedness (one- vs. two-sided) to alleviate alternative explanations that cannot be ruled out in Study 1. In Study 3, we varied the level of review two-sidedness (low vs. high two-sidedness) rather than its presence, and explored the underlying mechanisms. In Study 4A and Study 4B, we focused on one-sided reviews, examined the mechanisms in this situation, and also explored whether our findings are applicable for both positive and negative reviews. We also conducted a supplementary experiment to rule

out additional alternative explanations that might explain the findings of our main experiments.

2.3 Study 1

In the first study, we used real-world online reviews of mobile apps from Apple's App Store to test the hypotheses, because the review system in App Store represents a natural persuasion context where our independent and dependent variables can be quantified. Apple's App Store allows users to browse and purchase/download mobile applications (called "apps") designed for use on mobile devices. Existing users of an app can rate the app along a scale of 1 to 5 stars and submit a text review. In the text review, reviewers could describe their experience with the app using first-person and/or second-person pronouns, making it well suited to our investigation. Consumers who are interested in the app can read reviews from its former users before they make a purchase or download decision. They can also vote whether the review is helpful or not by clicking 'Yes' or 'No' buttons below the review. Both the number of 'Yes' votes and the total number of votes are displayed for reviews that have received at least one vote.

2.3.1 Data

We collected the data in April 2010 by first identifying apps ranked in the top 500 by popularity under each of the 20 categories (games, business, etc.) in the first three months of 2010. Among these apps, 40,417 had at least one review and we collected all their historical reviews. For each review, we recorded its rating, text review content, helpful votes, and total votes. We also recorded the following app-level information: average rating, count of all ratings, app category and whether or not the app was paid. In the end,

23

we collected 1,721,093 reviews in total. After filtering out reviews that were not written in English, had no content, or had a rating score of zero (presumably due to system errors), 1,623,497 reviews remained. Among this set, 418,415 reviews had received at least one vote.

2.3.2 Variables

We measured our dependent variable, review helpfulness, using the ratio of the number of helpful votes divided by the total number of votes (Mudambi and Schuff 2010; Yin et al. 2014). Therefore, our dependent variable was a proportion bounded between 0 and 1.

Our independent variable is operationalized based on reviewers' usage pattern of personal pronouns. Following prior research (Simmons et al. 2005), we used the text analysis software Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al. 2007), and calculated the ratio of second-person pronouns divided by the sum of first-person and second-person pronouns in each review. It is worth noting that some reviews may not contain either type of personal pronouns, such as short reviews like "This app is terrible." Because our independent variable cannot be quantified in such reviews, we excluded them and retained the rest (N = 301,517 reviews, around 72%) for the analyses.

We controlled for several variables that influence review helpfulness, including review rating, length, and reading difficulty (Korfiatis et al. 2008; Mudambi and Schuff 2010). Review length was measured by the number of words in a review. Reading difficulty was measured by the Gunning Fox Index (GFI), an estimate of the number of education years a student needs to understand a given text sample (Gunning 1969). We also controlled

for app-level variables, including average rating (of the app), number of ratings (the app has), whether or not the app is paid (coded 1 if paid, 0 otherwise), and app category. Summary statistics and correlations for these variables are presented in Tables 2-1 and 2-2.

Table 2-1 Summary Statistics.

Variable	N	Mean	Std. Dev.	Min	Max
Review Helpfulness	418415	0.59	0.42	0	1
Rating	418415	3.45	1.68	1	5
Length	418415	41.63	48.96	1	1134
Reading Difficulty	418415	7.05	4.13	0.4	461.6
Average Rating	418415	3.61	0.77	1	5
Count of Ratings	418415	499.27	709.19	1	3165
Paid	418415	0.57	0.49	0	1
Ratio of Second-Person Pronouns	301517	0.26	0.36	0	1

Table 2-2 Correlations (N=301,517).

	1	2	3	4	5	6	7	8
1. Review Helpfulness	1							
2. Ratio of Second-Person Pronouns	0.002	1						
3. Rating	0.366	-0.032	1					
4. Length	0.123	0.060	0.027	1				
5. Reading Difficulty	0.078	0.036	0.039	0.312	1			
6. Average Rating	0.104	-0.037	0.383	0.040	0.046	1		
7. Count of Ratings	-0.109	-0.006	-0.027	-0.062	-0.050	0.109	1	
8. Paid	0.060	-0.011	0.077	0.114	0.051	0.163	-0.084	1

2.3.3   Data Analysis and Results

Because the dependent variable, review helpfulness, was a proportion bounded between 0 and 1, OLS regression models may yield biased coefficients (Angrist and Pischke 2008; Kronmal 1993). To accommodate the bounded nature of this outcome variable and avoid the truncation problem that would result from dropping cases with 0 or 1 values, we adopted the fractional logit model that makes use of the logit link function and the binomial distribution of the dependent variable as our main analysis (see Baum 2008 for further discussions). The results are presented in Table 2-3. As shown in the main analysis (see Model 1), the coefficient for the ratio of second-person pronouns was positive and significant ($\beta = 0.036$, $p < 0.01$), providing initial evidence for our first hypothesis. Next, we calculated marginal effects that would be valuable for interpreting results of very large samples (Lin et al. 2013). The average marginal effect of this ratio was 0.008 ($p < 0.01$): as the ratio of second-person pronouns increases from 0% to 100%, review helpfulness increases by 0.8% (see also Papke and Wooldridge 1993; Wooldridge 2011). Thus, a subtle change in "invisible" personal pronouns can result in a small but significant change in review helpfulness perceptions.[3] In addition, the results showed that consistent with previous research (e.g., Korfiatis et al. 2008; Mudambi and Schuff 2010), review rating, length, and reading difficulty had significant and positive effects on review helpfulness.

---

[3] Although the effect size of personal pronouns appears small, note that this effect was caused by a subtle change in reviewers' use of personal pronouns ("I" vs. "you"). These personal pronouns do not convey substantive meanings, and they are largely "invisible" and "forgettable". As a result, our demonstration of a significant effect of personal pronouns in the archival study (despite its relatively small effect size) still suggests that these invisible function words should not be ignored.

Table 2-3 Empirical Results.

Variables	Model 1	Model 2	Model 3
	Fractional Logit Model (DV: review helpfulness)	2nd Stage of Heckman Model (DV: review helpfulness)	Negative Binomial Model (DV: number of helpful votes)
Number of Total Votes			0.105*** (0.001)
Rating	0.406*** (0.002)	0.089*** (0.000)	0.150*** (0.001)
Length	0.004*** (0.000)	0.001*** (0.000)	0.002*** (0.000)
Reading Difficulty	0.013*** (0.001)	0.003*** (0.000)	0.009*** (0.001)
Average Rating	-0.106*** (0.005)	-0.024*** (0.001)	-0.088*** (0.002)
Count of Ratings	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Paid	0.063*** (0.007)	0.031*** (0.002)	0.097*** (0.003)
Ratio of Second-Person Pronouns	0.036*** (0.009)	0.010*** (0.002)	0.027*** (0.004)
Category Dummies	Included	Included	
Constant	-1.042*** (0.037)	0.193*** (0.008)	-0.286*** (0.010)
N	301517	1061680	301517
Log Likelihood	-163583.96	-680586.08	-512614.33
Chi Square	46756.43	61532.97	47095.34

Notes: Robust standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

How do you begin a new chapter using this template? What must you do to get the page numbers to act correctly? Below are the steps for making a new chapter.

As a robustness check, to address a potential selection bias (i.e., not all reviews received votes), we employed Heckman's (1979) two-step sample selection model, with the first stage predicting the likelihood of a review being voted and the second stage predicting review helpfulness. Results of this analysis (see Model 2) were in line with results of our main analysis.

In addition, since the ratio of helpful votes might conceal the actual numbers of helpful votes or total votes (e.g., "1 out of 2 reviews is helpful" is equivalent to "50 out of 100 reviews are helpful" using the ratio measure), we conducted another robustness check with the total number of helpful votes included as an alternative measure of review helpfulness, and the total number of votes included as a covariate (Yin et al. 2017). Because the dependent variable, the number of helpful votes, was a count variable with its variance (50.05) greater than mean (2.36), we used negative binomial regression in this analysis (Chen and Lurie 2013; Yin et al. 2017). Consistent with our main analysis, results in Model 3 also revealed a positive and significant coefficient for the ratio of second-person pronouns ($\beta = 0.027$, $p < 0.01$).[4]

Finally, we conducted two additional robustness checks to account for differences between apps, because an app can have multiple reviews in our sample and our app-level control variables (e.g., average rating) may not fully capture the unobserved app-level

---

[4] Unlike Model 1 and Model 2, we did not include dummy variables for the app category, because the model cannot converge when category dummies were included.

heterogeneity. First, we utilized multilevel mixed-effects fractional logit model including both fixed- and random-effects. Specifically, we considered the independent variable, review-level control variables and app-level control variables from Model 1 as the fixed portion capturing fixed-effects analogous to standard regression coefficients. In addition, to specify the random-effects at the app level, we included an app-level random intercept capturing unobserved app-level heterogeneity that are not captured by app-level covariates in the model. Second, we employed multilevel mixed-effects negative binomial regression by including the fixed-portion from Model 3 along with an app-level random intercept. Results of both robustness checks (see Model 4 and Model 5 in Table 2-4) were consistent with the main analyses.

Table 2-4 Results of Additional Robustness Checks.

Variables	Model 4	Model 5
	Fractional Logit Model (DV: review helpfulness)	Negative Binomial Model (DV: number of helpful votes)
Number of Total Votes		0.099***
		(0.004)
Rating	0.424***	0.150***
	(0.011)	(0.004)
Length	0.007***	0.002***
	(0.000)	(0.000)
Reading Difficulty	0.013***	0.007***
	(0.002)	(0.001)
Average Rating	-0.154***	-0.110***
	(0.015)	(0.005)
Count of Ratings	-0.000***	-0.000***
	(0.000)	(0.000)
Paid	0.246***	0.109***
	(0.019)	(0.006)
Ratio of Second-Person Pronouns	0.076***	0.040***
	(0.014)	(0.005)
Category Dummies	Included	
Constant	-0.060	-0.222***
	(0.105)	(0.021)
N	301517	301517
Log Likelihood	-143887.57	-509059.82
Chi Square	4683.04	9648.66

Notes: Robust standard errors in parentheses; Variance of random intercepts omitted; * p<0.1, ** p<0.05, *** p<0.01.

2.3.4   Discussion

In this study, we tested the main effect of reviewers' use of personal pronouns on review helpfulness using actual reviews from Apple's App Store. The results showed that the percentage of second-person pronouns was positively associated with review helpfulness, providing real-world evidence for H1.

However, the use of archival data necessitated a major limitation: it could not provide direct insights for the causal impact of a reviewer's use of personal pronouns on review helpfulness. Although we controlled for a variety of variables that have been shown to influence review helpfulness, unobserved factors that correlate with reviewers' use of personal pronouns and also influence readers' perception of review helpfulness present possibilities for alternative explanations. In particular, the influence of reviewers' use of personal pronouns may not arise from personal pronouns per se, but from differences in review content that are driven by differential efforts of reviewers who prefer to use second-person pronouns versus those who prefer to use first-person pronouns. For example, reviewers who prefer the use of second-person pronouns may at the same time provide more suggestions and opinions to future readers. In addition, the archival data does not allow us to precisely measure the two-sidedness of review arguments, which is necessary for testing our second hypothesis. We designed an experiment in the next study to address these issues.

2.4   Study 2

The primary goals of Study 2 were to isolate the causal impact of personal pronouns on review helpfulness and explore the moderating role of review two-sidedness.

31

Participants took part in a hypothetical online decision-making task in which they read reviews of four different mobile apps and then evaluated the reviews and reviewers. We manipulated personal pronouns within-subjects at two levels (first-person only vs. second-person only) and review two-sidedness between-subjects at two levels (one-sidedness vs. two-sidedness).

2.4.1 Stimulus Materials

We utilized a time management app in the experiment because time management is a universal concern most people are familiar with. Time management apps can improve users' productivity and help them beat procrastination through a time management method in which a timer is used to break down working time into intervals, separated by short breaks.

We developed two sets of treatment reviews for the one-sided condition in two steps. First, we consulted actual reviews from Apple's App Store and created two positive reviews containing first-person pronouns. The two reviews described different user experiences using only first-person pronouns (e.g., "I," "me," "mine"). We fixed the valence of reviews to be positive in order to remove the influence of valence. In the second step, we constructed two corresponding treatment reviews that replaced first-person pronouns with second-person pronouns. Within each set of reviews, the only difference between the two versions is personal pronouns.

We then developed two sets of treatment reviews for the two-sided condition. For each positive review (containing first- or second-person pronouns) created in one-sided condition, we constructed a corresponding two-sided review by changing 2 (out of 4)

statements contained in the review to be negative in valence (e.g., using antonyms and adding negations) while holding the substantive content identical. We also kept the number of words in each review at a similar level (around 50); the only difference between the one- and two-sided conditions was review two-sidedness. We also added "Pros" and "Cons" in each review to strengthen the manipulation of review two-sidedness. All treatment reviews are presented in Table 2-5.

Table 2-5 Review Stimuli in Study 2.

	#	Review Written in First-Person Pronouns	Review Written in Second-Person Pronouns
One-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. • It keeps the number of sessions **I** have accomplished and categorizes **my** tasks. • It's easy to organize **my** own time because **I** can customize the timer. **Cons:** None.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. • It keeps the number of sessions **you** have accomplished and categorizes **your** tasks. • It's easy to organize **your** own time because **you** can customize the timer. **Cons:** None.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. • The ticking sound is clear, so **I** can remember when **I** should take a rest. • In addition, it can sync **my** tracked data between **my** phone and computer. **Cons:** None.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. • The ticking sound is clear, so **you** can remember when **you** should take a rest. • In addition, it can sync **your** tracked data between **your** phone and computer. **Cons:** None.
Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **I** have accomplished or categorize **my** tasks. • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **you** have accomplished or categorize **your** tasks. • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **I** *cannot* remember when **I** should take a rest. • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **you** *cannot* remember when **you** should take a rest. • It *cannot* sync **your** tracked data between **your** phone and computer.

Notes: emphases and italics added for illustration purpose only (i.e., not shown to participants).

34

2.4.2 Procedure and Measures

159 respondents (75 male) from Amazon MTurk participated in this study and were compensated for their participation. 98 percent of them were originally from the United States, 64 percent achieved bachelor's degree or above as their highest education level, and their average age was 49.

In the cover story, participants were asked to imagine that they were looking for a time management app from Apple's App Store, and their search returned three apps. These three apps had similar average rating score of 4 (out of 5) stars, and each costed $2. The participants were told to evaluate reviews of three apps before their final decision, and they were shown three reviews (each from a different app), one at a time. For each app, they were asked to read an online review randomly selected from its former users. One "filler" review without any personal pronouns was presented in position 1. The two treatment reviews (one written in first-person pronouns and the other written in second-person pronouns) chosen from different sets were presented in positions 2 and 3, with the order of treatment counterbalanced. Each participant was randomly assigned to the one-sided or two-sided condition.

After reading each review, participants were asked to report their perceptions of review helpfulness using a 9-point scale. We utilized a measure of perceived review helpfulness with three items adapted from Sen and Lerman (2007) (e.g., "not at all helpful/very helpful"). As a manipulation check, we also asked participants to report their perception of review two-sidedness using three items adapted from Jensen et al. (2013) (e.g., "very one-sided/very two-sided"). To check whether our manipulation of personal

pronouns was successful, we conducted a separate pretest (see details below). Appendix A contains all the measures used in this experiment and follow-up studies.

2.4.3 Results

Before further analysis, we conducted manipulation checks to ensure that our manipulations of review two-sidedness and personal pronouns were successful. We first utilized ANCOVA to check the manipulation of review two-sidedness, with perceived two-sidedness as the dependent variable, personal pronouns entered as a within-subjects factor, two-sidedness manipulation entered as between-subjects factor, and treatment order entered as a covariate. Results showed that perceived two-sidedness in the one-sided condition was significantly lower than that in the two-sided condition ($M = 3.37$ vs. 5.78, $F(1, 156) = 111.14$, $p < 0.001$). In addition, to check whether our manipulation of personal pronouns was successful, we recruited a separate group of 73 subjects from Amazon MTurk. The cover story and the procedure were similar to the main study, except that participants were only asked to report their perception of reviewers' use of first-person and second-person pronouns in each treatment review using two items adapted from Pennebaker et al. (2003) and Simmons et al. (2005) (see Appendix A for the measure). A similar ANCOVA, with perceived frequency of personal pronouns included as the dependent variable, revealed that the review written in first-person pronouns was perceived to contain significantly more first-personal pronouns ($M = 7.10$ vs. 2.89, $F(1, 70) = 106.29$, $p < 0.001$) and significantly less second-person pronouns ($M = 2.91$ vs. 7.21, $F(1, 70) = 108.59$, $p < 0.001$) than the review written in second-person pronouns. Thus, the manipulations of both our independent and moderating variables were successful.

Next, we utilized a similar ANCOVA to test the effect of personal pronouns on perceived review helpfulness and the moderating effect of review two-sidedness. Results revealed that the overall effect of second-person (vs. first-person) pronouns on review helpfulness did not reach significance ($M = 6.92$ vs. 7.08, $F(1, 156) = 1.32$, $p = 0.252$). However, consistent with H2 and Study 1, the interaction between personal pronouns and review two-sidedness was marginally significant ($F(1, 156) = 3.37$, $p = 0.068$). Within the one-sided condition, the pairwise comparison showed that perceived helpfulness did not significantly differ between reviews written in first-person pronouns and reviews written in second-person pronouns ($M = 6.87$ vs. 6.76, $F(1, 156) = 0.23$, $p = 0.636$); within the two-sided condition, the pairwise comparison revealed a significant increase in perceived helpfulness from the reviews containing first-person pronouns to the reviews containing second-person pronouns ($M = 6.99$ vs. 7.40, $F(1, 156) = 4.67$, $p = 0.032$). Thus, we obtained no evidence for H1 but initial evidence for H2 in this study (see Figure 2-2).
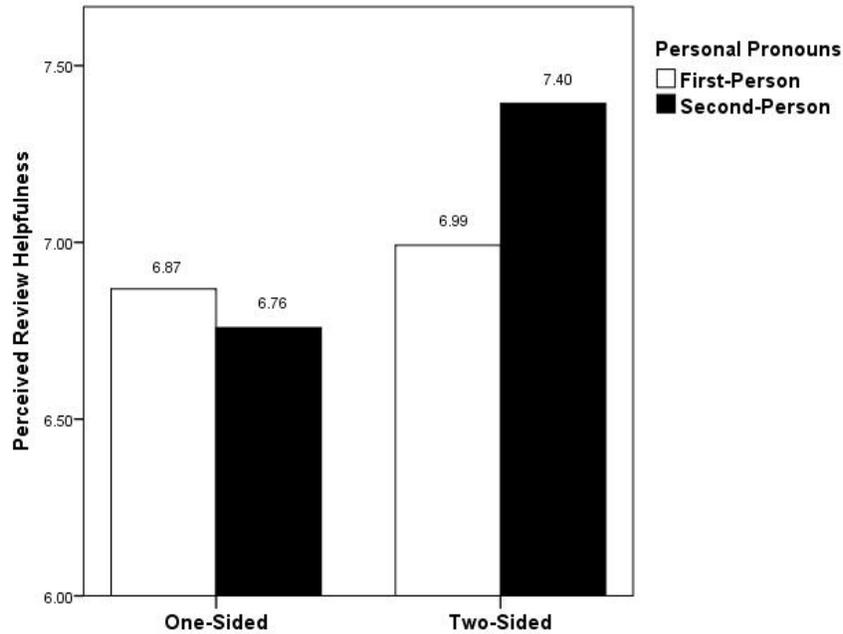
Figure 2-2 Main Results in Study 2.

2.4.4 Discussion

While Study 1 found an overall positive effect of personal pronouns on review helpfulness as hypothesized in H1, the current study failed to replicate this finding. A possible reason is that reviewers who prefer to use second-person pronouns may also write the substantive content of reviews in different ways than those who prefer to use first-person pronouns. On the other hand, this study provided initial evidence for the moderator we proposed in H2: the positive effect of second-person (vs. first-person) pronouns on perceived helpfulness was much stronger for two-sided reviews while this positive effect disappeared for one-sided reviews.

This study had a number of limitations that were addressed in follow-up experiments. First, there could be different (e.g., low and high) levels of two-sidedness in

an actual review. Thus, we designed the next experiment to explore whether the positive effect of second-person (vs. first-person) pronouns for a two-sided review would remain when the review was less two-sided. Second, neither of the first two studies examined the mediating processes that might underlie the effect of second-person (vs. first-person) pronouns on perceived review helpfulness. In the following experiments, we included measures of perceived empathic concern and perceived persuasion motives to explore the probable underlying processes.

2.5 Study 3

In Study 3, we examined whether the positive effect of second-person (vs. first-person) pronouns that we observed in the two-sided condition of the previous study would remain significant under different levels of two-sidedness, and also explored the underlying processes. We varied personal pronouns in the same manner as in Study 2, but manipulated review two-sidedness more subtly through its extent rather than its presence.

2.5.1 Stimulus Materials

We developed stimuli for this study based on the treatment reviews from Study 2. First, we kept two sets of reviews used in the two-sided condition of Study 2 as "high two-sided" condition in the current study, because they contain an equal number of positive and negative statements. Second, we created low two-sided reviews that are slightly positive because the majority of reviews in the real world are positive, and that positive reviews discussing something negative are more commonly observed than negative reviews discussing something positive. Specifically, we constructed two corresponding sets of reviews by decreasing the number of negative statements from 2 to

39

1 (out of 4) and accordingly increasing the number of positive statements from 2 to 3. Again, we held the total amount of information and substantive content identical; the only difference between the low and high two-sided conditions was the extent of two-sidedness. All treatment reviews are present in Table 2-6.

Table 2-6 Review Stimuli in Study 3.

	#	Review Written in First-Person Pronouns	Review Written in Second-Person Pronouns
Low Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. • It keeps the number of sessions **I** have accomplished and categorizes **my** tasks. **Cons:** • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. • It keeps the number of sessions **you** have accomplished and categorizes **your** tasks. **Cons:** • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. • The ticking sound is clear, so **I** can remember when **I** should take a rest. **Cons:** • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. • The ticking sound is clear, so **you** can remember when **you** should take a rest. **Cons:** • It *cannot* sync **your** tracked data between **your** phone and computer.
High Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **I** have accomplished or categorize **my** tasks. • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **you** have accomplished or categorize **your** tasks. • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **I** *cannot* remember when **I** should take a rest. • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **you** *cannot* remember when **you** should take a rest. • It *cannot* sync **your** tracked data between **your** phone and computer.

Notes: emphases and italics added for illustration purpose only (i.e., not shown to participants).

41

2.5.2 Procedure and Measures

191 undergraduate students (71 male) from a U.S. university participated in this experiment in exchange for extra credit. 96 percent were originally from the United States, 91 percent were juniors or above, and the average age of the students was 21.

This study followed the similar procedure as in Study 2, except that we also asked participants to report their perceptions of the reviewer's empathic concern and the reviewer's persuasion motives along 9-point scales. After reading each review, participants were asked to report their perceptions of review helpfulness using the same items as in Study 2. We then asked participants to report their perception of the reviewer's empathic concern using three items adapted from Goldstein et al. (2014) and Toi and Batson (1982) (e.g., "When reading about this review, to what extent do you think this reviewer understands your feelings?"), and their perception of the reviewer's persuasion motives using four items adapted from Campbell and Kirmani (2000) and Williams et al. (2004) (e.g., "While I was reading the review, I thought it was pretty obvious that the reviewer was trying to influence me."). Participants also rated the level of review two-sidedness using the same items as in Study 2. See Appendix A for all the measures used in this study.

2.5.3 Results

First, we conducted a manipulation check of review two-sidedness. Results revealed that perceived two-sidedness in the low two-sidedness condition was significantly lower than that in the high two-sidedness condition ($M = 5.22$ vs. 5.95, $F(1, 188) = 18.07$, $p < 0.001$). Thus, the manipulation of review two-sidedness was successful.

42

Next, we conducted ANCOVA to explore the effect of personal pronouns on perceived review helpfulness under different levels of review two-sidedness, with second-person vs. first-person pronouns entered as a within-subjects factor, level of two-sidedness entered as a between-subjects factor, and treatment order entered as a covariate. Results revealed that the interaction between the effect of personal pronouns and level of review two-sidedness did not reach significance ($F(1, 188) = 0.10$, $p = 0.753$). In addition, reviews written in first-person pronouns were perceived to be less helpful than reviews written in second-person pronouns ($M = 6.71$ vs. 7.17, $F(1, 188) = 11.09$, $p = 0.001$). Pairwise comparisons revealed that reviews with first-person pronouns were considered less helpful than reviews with second-person pronouns when two-sidedness was low ($M = 6.75$ vs. 7.26, $F(1, 188) = 6.62$, $p = 0.011$), as well as when two-sidedness was high ($M = 6.66$ vs. 7.08, $F(1, 188) = 4.56$, $p = 0.034$). Thus, we obtained consistent evidence for the effect of personal pronouns on perceived helpfulness when the two-sidedness of the review is low (see Figure 2-3).
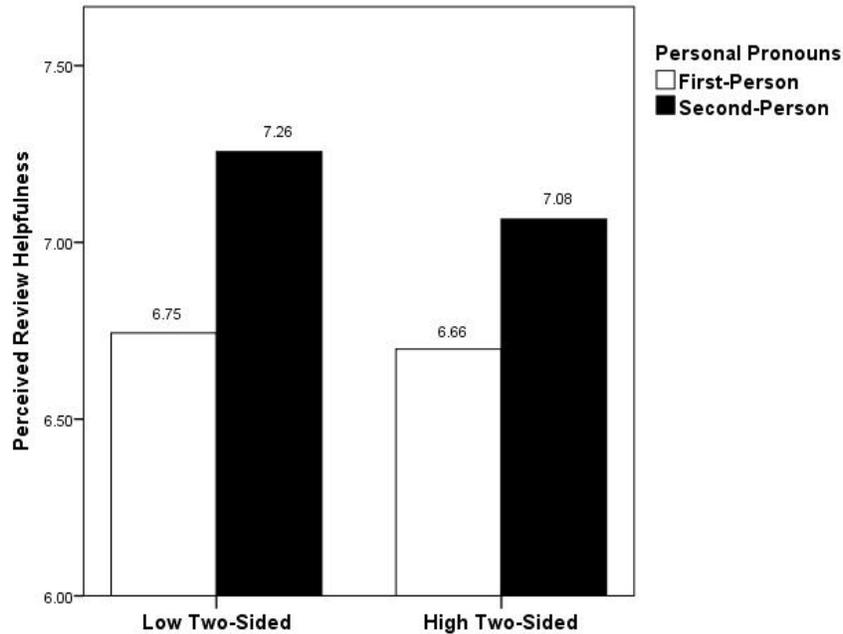
Figure 2-3 Main Results in Study 3.

In addition, we examined the probable mechanisms—perceived empathic concern and perceived persuasion motives—that could underlie the effect of personal pronouns. We first conducted ANCOVA to examine the effects of second-person (vs. first-person) pronouns on these two mediators. Results revealed that second-person (vs. first-person) pronouns significantly increased perceived empathic concern ($M = 5.37$ vs. 5.97, $F(1, 188) = 17.24$, $p < 0.001$), but they did not significantly increase perceived persuasion motives ($F(1, 188) = 0.42$, $p = 0.516$).

Then we conducted a formal mediation analysis for each two-sided condition based on bootstrapping, using SPSS macro MEMORE developed by Montoya and Hayes (2017). Compared to the conventional approach proposed by Judd el al. (2001), the bootstrapping method eliminates the requirement of discrete hypothesis tests and allows us to test multiple mediators in parallel in a within-subjects design (Montoya and Hayes

44

2017). Results under low two-sided condition were consistent with the ANCOVA results above: a reviewer's use of second-person (vs. first-person) pronouns had a positive effect on perceived empathic concern ($\beta = 0.57$, $t(95) = 2.83$, $p = 0.006$), which in turn had a positive effect on perceived review helpfulness ($\beta = 0.60$, $t(91) = 7.37$, $p < 0.001$). The indirect effect of personal pronouns on review helpfulness through perceived empathic concern was positive and significant as zero was not included in its bias-corrected confidence interval ($\beta = 0.34$, 95% CI = [0.10, 0.62]). On the other hand, the effect of a reviewer's use of second-person pronouns on perceived persuasion motives did not reach significance ($\beta = 0.03$, $t(95) = 0.33$, $p = 0.739$). The indirect effect through perceived persuasion motives was also insignificant ($\beta = 0.01$, 95% CI = [-0.04, 0.07]) (see Figure 2-4). Results of MEMORE analyses under high two-sided condition were consistent with those under low two-sided condition. Taken together, the positive process (i.e., perceived empathic concern) overweighed the negative process (i.e., perceived persuasion motives) in two-sided reviews, resulting in an overall positive effect of second-person (vs. first-person) pronouns on perceived review helpfulness.
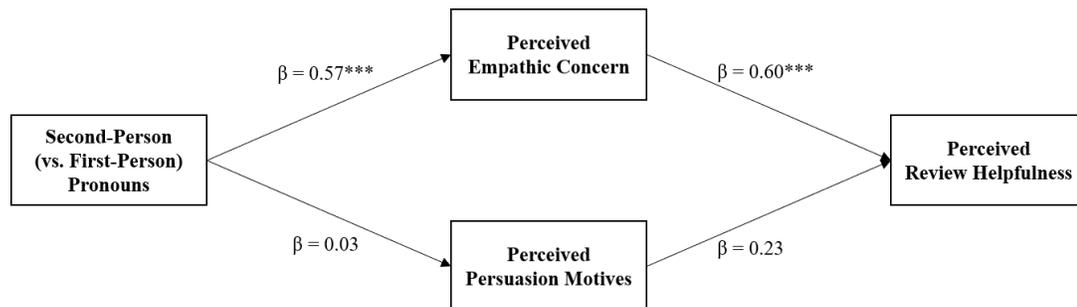
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 2-4 Mediation Results in Study 3.

2.5.4   Discussion

This study provided evidence that the positive effect of second-person (vs. first-person) pronouns on perceived review helpfulness could exhibit not only for a highly two-sided review, but also for a low two-sided review. We also explored the mechanisms underlying the effect of personal pronouns in two-sided reviews, finding evidence for the positive mechanism of perceived empathic concern, but not the negative mechanism.

Although we revealed the mediating role of empathic concern in driving the positive effect of second-person (vs. first-person) pronouns in two-sidedness reviews, it remains unclear what processes drive the null effect of personal pronouns in one-sided reviews. One possible reason is that a negative process could be activated under one-sidedness condition, and that this negative process cancels out the positive process of empathic concern. In the next two experiments, we focused on one-sided reviews and explored this possibility.

2.6   Study 4A

In the next two studies, we examined the probable mechanisms underlying the main effect of second-person (vs. first-person) pronouns on perceived helpfulness of one-sided reviews. In Study 4A, we fixed the valence of treatment reviews to be positive to remove the influence of valence. We created two sets of reviews based on the review stimuli used in the one-sidedness condition of Study 2. Because the valence of reviews was fixed to be positive and all the treatment reviews were one-sided (i.e., including only positive statements), we also removed terms "Pros" and "Cons" and bullet points in each

46

treatment review (see Table 2-7). We manipulated personal pronouns within-subjects at two levels, similarly as in earlier experiments.

Table 2-7 Review Stimuli in Study 4A

#	Review Written in First-Person Pronouns	Review Written in Second-Person Pronouns
1	This well-designed application helps me establish a high level of focus on my work and study. It's surprising that I can observe, in several forms, how I have been doing to meet my goals each week. I like this app.	This well-designed application helps you establish a high level of focus on your work and study. It's surprising that you can observe, in several forms, how you have been doing to meet your goals each week. You'll like this app.
2	I love this powerful app because it allows me to save and track my progress over time. It keeps a record of my work time as well. After a certain period of tracking data, I can better understand my productivity pattern.	You'll love this powerful app because it allows you to save and track your progress over time. It keeps a record of your work time as well. After a certain period of tracking data, you can better understand your productivity pattern.

Notes: emphases added for illustration purpose only (i.e., not shown to participants).

2.6.1 *Procedure and Measures*

88 respondents (33 male) from Amazon MTurk participated in this study and were compensated for their participation. 95 percent were originally from the United States, 88 percent achieved bachelor's degree or above as the highest education level, and the average age was 37.

The cover story and procedure were similar as in Study 3, except that we dropped the filler review to reduce the number of reviews each participant was asked to read and evaluate. Therefore, each participant was told to evaluate reviews of two apps before their final decision, and they were shown two reviews (each from a different app), one at

a time. The two treatment reviews (one with first-person pronouns and the other with second-person pronouns) were chosen from different sets and the order of treatment was counterbalanced. Similar as in Study 3, after reading each review, participants were asked to report their perceptions of review helpfulness, the reviewer's empathic concern, and the reviewer's persuasion motives using a 9-point scale. All the measures were the same as in previous studies.

2.6.2   Results

We utilized ANCOVA to explore the effect of personal pronouns on perceived review helpfulness, with personal pronouns entered as a within-subjects factor and treatment order entered as a covariate. Results revealed that the effect of second-person (vs. first-person) pronouns on review helpfulness did not reach significance ($M = 7.01$ vs. 6.98, $F(1, 86) = 0.02$, $p = 0.898$). Thus, we did not obtain evidence for H1 in this study.

Despite the lack of a main effect, further mediation tests are still valid and necessary if opposing mediating processes might exist (Hayes 2009; Hayes and Rockwood 2017; MacKinnon et al. 2000). When indirect effects operate in opposite directions and cancel each other out, the main effect could manifest as being insignificant. To examine this possibility, we first used ANCOVA to examine the effects of personal pronouns on the two mediators. Results revealed that second-person (vs. first-person) pronouns significantly increased perceived empathic concern ($M = 5.08$ vs. 5.60, $F(1, 86) = 4.11$, $p = 0.046$) and perceived persuasion motives ($M = 3.26$ vs. 3.92, $F(1, 86) = 15.27$, $p < 0.001$).

Then we conducted a formal mediation analysis using SPSS macro MEMORE

(Montoya and Hayes 2017). Results revealed that a reviewer's use of second-person (vs.

first-person) pronouns had a positive effect on perceived empathic concern ($\beta = 0.50$,

$t(87) = 1.92$, $p = 0.058$), which in turn had a positive effect on perceived review

helpfulness ($\beta = 0.47$, $t(83) = 6.42$, $p < 0.001$). The indirect effect of personal pronouns

through perceived empathic concern was positive and significant ($\beta = 0.23$, 95% CI =

[0.01, 0.49]). In addition, a reviewer's use of second-person (vs. first-person) pronouns

also had a positive effect on perceived persuasion motives ($\beta = 0.67$, $t(87) = 3.90$, $p <$

0.001), which then had a negative effect on perceived review helpfulness ($\beta = -0.30$, $t(83)$

$= -2.72$, $p = 0.008$). The indirect effect through perceived persuasion motives was also

significant ($\beta = -0.20$, 95% CI = [-0.41, -0.03]) (see Figure 2-5). Taken together, a lack of

main effect of personal pronouns in one-sided positive reviews could be caused by two

mediators operating in opposite directions and canceling each other out.
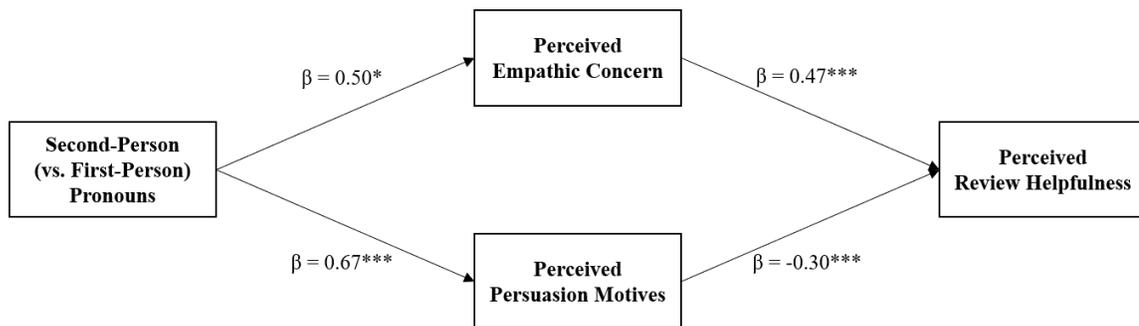
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 2-5 Mediation Results in Study 4A.

2.6.3 Discussion

This experiment provided evidence for the co-existence of two opposing processes for one-sided positive reviews, which can help explain the lack of a main effect of personal pronouns on perceived review helpfulness. Because all previous studies focused on one-sided *positive* reviews or two-sided reviews, we conducted a final study to examine the effect of personal pronouns for one-sided *negative* reviews.

2.7 Study 4B

In this study, we focused on one-sided *negative* reviews and tested the effect of personal pronouns on perceived review helpfulness and its underlying mechanisms. Specifically, for each positive review used in Study 4A, we constructed a corresponding negative review by changing all statements to negative valence (see Table 2-8). 58 undergraduate students (24 male) from a U.S. university participated for exchange of extra credit. 86 percent were originally from the United States, 50 percent were juniors or above, and the average age was 20. The cover story, the procedure, and all measures used in this study were similar as in Study 4A.

Table 2-8 Review Stimuli in Study 4B.

#	Review Written in First-Person Pronouns	Review Written in Second-Person Pronouns
1	This badly-designed application doesn't help me establish a high level of focus on my work and study. It's surprising that I cannot observe, in any form, how I have been doing to meet my goals each week. I don't like this app.	This badly-designed application doesn't help you establish a high level of focus on your work and study. It's surprising that you cannot observe, in any form, how you have been doing to meet your goals each week. You won't like this app.
2	I don't love this powerless app because it doesn't allow me to save and track my progress over time. It doesn't keep a record of my work time as well. After a certain period of tracking data, I cannot better understand my productivity pattern.	You won't love this powerless app because it doesn't allow you to save and track your progress over time. It doesn't keep a record of your work time as well. After a certain period of tracking data, you cannot better understand your productivity pattern.

Notes: emphases added for illustration purpose only (i.e., not shown to participants).

2.7.1   Results

We utilized ANCOVA to explore the effect of personal pronouns on perceived review helpfulness. Results revealed that the effect of second-person pronouns did not reach significance ($M = 6.59$ vs. 6.43, $F(1, 56) = 0.31$, $p = 0.580$), similar as what we observed in the previous study.

We then examined the mediation effects of perceived empathic concern and perceived persuasion motives in two steps. ANCOVA results revealed that a reviewer's use of second-person (vs. first-person) pronouns had a positive effect on perceived empathic concern ($M = 4.55$ vs. 5.60, $F(1, 56) = 12.11$, $p = 0.001$) and perceived persuasion motives ($M = 3.57$ vs. 4.26, $F(1, 56) = 6.27$, $p = 0.015$). Second, we used SPSS macro MEMORE to conduct a formal mediation analysis. Results revealed that

second-person (vs. first-person) pronouns had a positive effect on perceived empathic concern (β = 1.05, $t(57)$ = 3.26, p = 0.002), which in turn had a positive effect on perceived review helpfulness (β = 0.66, $t(53)$ = 6.14, $p < 0.001$). The indirect effect of personal pronouns through perceived empathic concern was positive and significant (β = 0.70, 95% CI = [0.27, 1.17]). In addition, second-person (vs. first-person) pronouns had a positive effect on perceived persuasion motives (β = 0.69, $t(57)$ = 2.20, p = 0.032), which in turn had a negative effect on perceived review helpfulness (β = -0.35, $t(53)$ = -3.16, p = 0.003). The indirect effect through perceived persuasion motives was also significant (β = -0.24, 95% CI = [-0.60, -0.01]) (see Figure 2-6).
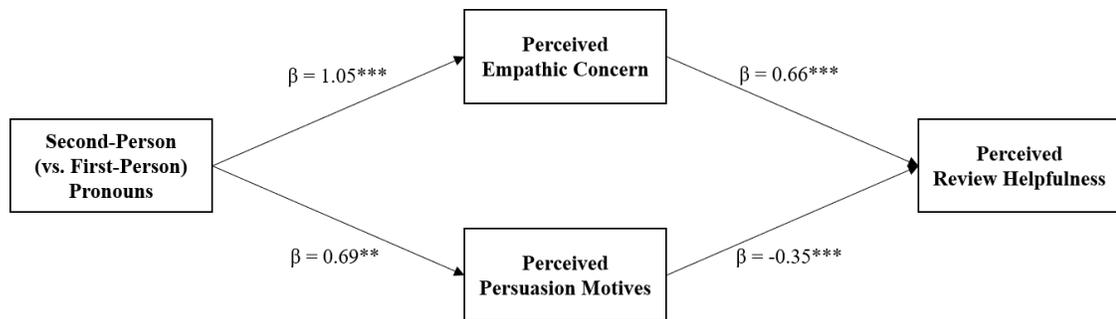
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 2-6 Mediation Results in Study 4B.

2.7.2 Discussion

In Study 4A and 4B, we replicated the earlier finding of an insignificant main effect of personal pronouns on the perceived helpfulness of one-sided reviews, regardless of the valence of the reviews. In addition, both studies provided evidence that for one-sided reviews, the null effect of personal pronouns might be explained by opposing mediating processes operating in opposite directions and canceling each other out.

In a supplementary experiment (see Appendix B), we investigated whether the impact of personal pronouns on review helpfulness could be explained by perceived politeness, perceived psychological closeness, perceived interpersonal closeness, and perceived altruism. In addition, we explored whether our manipulation of review two-sidedness also varied review comprehensiveness. Results of this supplementary experiment ruled out perceived politeness, perceived psychological closeness, perceived interpersonal closeness, and review comprehensiveness as alternative explanations for the findings of our main experiments. On the other hand, results revealed that the effect of personal pronouns might be explained by readers' perception of reviewer altruism. However, this does not conflict with our empathy-related arguments because of their close association (e.g., Coke et al. 1978), such that the more second-person pronouns reviewers use in reviews, the more likely that the reviewers are perceived to be empathic and concerned about readers (and correspondingly more willing to help readers).

2.8   General Discussion

Utilizing distinct methodologies, the five studies provided converging evidence that reviewers' use of personal pronouns plays a nontrivial role in review readers' perception of review helpfulness, and that the effect of second-person (vs. first-person) pronouns is greater for two-sided reviews compared with one-sided reviews. The major findings are summarized in Table 2-9. For two-sided reviews, our results from Study 3 suggested that the positive process (i.e., perceived empathic concern) overweighs the negative process (i.e., perceived persuasion motives), resulting in a strongly positive effect of second-person (vs. first-person) pronouns. For one-sided reviews, however, our results from Study 4A and 4B showed that the positive process and the negative process

53

could operate in opposite directions, leading to a null effect of personal pronouns on review helpfulness.

Table 2-9 Summary of Findings.

	H1: A reviewer's use of second-person (vs. first-person) pronouns in a review is positively associated with the helpfulness of the review perceived by readers.	H2: The positive effect of a reviewer's use of second-person (vs. first-person) pronouns on perceived review helpfulness is greater for two-sided reviews compared with one-sided reviews.
Study 1	Overall Effect Supported	Not Explored
Study 2	Overall Effect Not Supported (we explored the reasons in the next 3 studies)	Supported
Study 3 (Two-Sidedness)	Supported (because positive mechanism overweighs negative mechanism in two-sided reviews)	Consistent with H2
Study 4A (One-Sided Positive)	Not Supported (because positive and negative mechanisms cancel out each other in one-sided positive reviews)	Consistent with H2
Study 4B (One-Sided Negative)	Not Supported (because positive and negative mechanisms cancel out each other in one-sided negative reviews)	Consistent with H2

2.8.1 Theoretical Implications

Our paper makes a number of unique theoretical contributions. First, our research contributes to the online review literature by going beyond content words that convey real meanings. Prior studies have investigated the influence of the characteristics of review text on review helpfulness (e.g., Jensen et al. 2013; Yin et al. 2014). However, those studies mainly identified the characteristics based on the content words without considering reviewers' use of function words. Although functions words such as personal pronouns do not carry real meanings and are largely "invisible" and "forgettable", they reflect people's personality and psychological states (Pennebaker 2011); thus, their psychological functions for individuals should not be ignored. We reveal the nontrivial impacts of reviewers' unconscious use of personal pronouns on consumers' helpfulness perception of reviews. Specifically, our findings indicate that personal pronouns can affect consumers' perception of review helpfulness through two opposing processes (i.e., perceived empathic concern and perceived persuasion motives), which in turn affect review helpfulness in opposite directions. Therefore, this novel angle extends the focus of prior research studying antecedents of review helpfulness, and points to reviewers' use of personal pronouns as another important driver of consumers' helpfulness perceptions.

Second, our findings also contribute to the linguistic literature by extending the role of personal pronouns from the offline setting to the online environment. On the one hand, the examination of second-person (vs. first-person) pronouns in the offline settings generally involves two individuals (e.g., Ickes et al. 1986; Simmons et al. 2005), but the use of personal pronouns in the online environment is different and might involve more than two individuals. In our context, reviewers who use more second-person (vs. first-

person) pronouns may craft reviews not for an individual reader, but for hundreds of thousands of prospective online consumers who may read the reviews and make purchase decisions. Research in offline settings found that a focus on others (vs. self) plays a significant role even between strangers in initial encounters (Fraley and Aron 2004; Galinsky and Moskowitz 2000). Our results reveal that a reviewer's use of personal pronouns also plays a significant role in influencing prospective consumers' (review readers') perceptions and inferences, even though these consumers do not have any interaction with the reviewer before. The findings extend the study of personal pronouns from the offline setting typically involving only two individuals to an online environment where a shift in personal pronouns can influence hundreds of thousands of prospective consumers.

On the other hand, we take advantage of the unique nature of online reviews and explore a boundary condition for the impact of personal pronouns on perceived review helpfulness. Because the use of second-person (vs. first-person) pronouns and other-focused attention have been revealed to bring various benefits in the offline context (e.g., Coke et al. 1978; Galinsky and Moskowitz 2000; Hodges et al. 2011), it is reasonable to expect the greater use of second-person (vs. first-person) pronouns to also lead to positive outcomes in the online environment. However, online reviews are unique in that their primary purposes are to inform and persuade future consumers (Sparks et al. 2013), while we know little about the role of personal pronouns in such persuasion contexts. By identifying review two-sidedness as a critical boundary condition and suggesting possible reasons, our findings reveal that reviewers' use of second-person (vs. first-person) pronouns may not always increase review helpfulness. In particular, the use of second-

person (vs. first-person) pronouns plays a significant role in two-sided reviews, while it does not influence the perceived helpfulness of one-sided reviews with the co-existence of two probable opposing processes (i.e. perceived empathic concern and perceived persuasive motives). By examining the contingency for the influence of personal pronouns in online reviews, this paper reveals that the benefit of personal pronouns may not always materialize in persuasion contexts, deepens our understanding of when and how the use of personal pronouns can be effective in influencing persuasion targets, and suggests the need of more research to explore the nuanced role of personal pronouns in online environments.

2.8.2   *Practical Implications*

Our findings offer practical implications for product/service providers. Because more helpful reviews are presumably more influential for consumers, product/service providers stand to benefit from identifying helpful reviews early on and dealing with them more proactively. We find that reviews that contain more second-person (vs. first-person) pronouns are perceived more helpful only if the reviews contain mixed opinions. As a result, companies can identify mixed-opinion reviews that contain more second-person (relative to first-person) pronouns, and address the negative opinions proactively to reduce their potential damages. For example, they may offer a genuine response directly addressing the reviewers' concerns.

In addition, our findings also provide insights for review platforms and reviewers. With more helpful reviews, a review platform can increase consumer perceptions of the platform and increase site "stickiness," encouraging consumers to re-visit or spend longer

time at the site (Kumar and Benbasat 2006; Mudambi and Schuff 2010). Based on our results, review platforms might consider incorporating the use of function words such as personal pronouns in their review-writing guidelines to encourage the creation of more helpful reviews. For example, the guideline may simply ask reviewers to shift their use of personal pronouns from first-person to second-person or to shift their focus of attention from self to others, assuming that reviewers would follow such guidelines. A caveat worth pointing out is that merely shifting personal pronouns from "I" to "you" in the reviews may not always increase review helpfulness, and this strategy may have unintended consequences. By the same token, reviewers striving to provide more helpful content should be aware that simply switching personal pronouns may not be sufficient especially if the opinions are purely positive or purely negative; instead, a shift in personal pronouns combined with corresponding changes to the substantive content (e.g., writing longer, in more depth, etc.) may be most effective in boosting review helpfulness. That being said, switching from "I" to "you" is a much simpler strategy of writing reviews than improving the quality or length of reviews that have been advocated in prior research, and it is advisable especially when reviewers have mixed opinions.

2.8.3 Limitations and Future Research

Our research also has a few limitations that provide avenues for future research. First, although we found the overall positive effect of other-focus on review helpfulness in Study 1, we did not obtain evidence for this overall positive effect in Study 2. One possible explanation for this inconclusive finding is the distinct nature of experiments compared with the first study. We varied personal pronouns (first-person only vs. second-person only) in the experiments while holding the review content identical, but

substantive content of reviews could co-vary with the use of second-person (vs. first-person) pronouns in the archival study. In the archival study, reviewers who focus on others may not only change "I" to "you", but also craft substantive content in different manners. Thus, the use of second-person (vs. first-person) pronouns and associated attentional focus may also influence the writing process of review writers and ultimately the characteristics of review text they produce (e.g., breadth of discussed topics, depth of opinions, etc.). The effect of personal pronouns and attentional focus on reviewer behavior is beyond the scope of the current paper, but it is surely worthy of future investigation.

Second, we varied the use of personal pronouns at two extreme levels (first-person only vs. second-person only) in the experiments. However, we did not consider the impact of other combinations of personal pronouns, such as a mix of first-person and second-person pronouns, third-person pronouns only, as well as the interplay of first-, second-, and third-person pronouns. Although the role of a mix of different types of personal pronouns in online reviews is out of this paper's scope, it is a very interesting direction for future research. For example, we did not explore the role of third-person pronouns in this paper because third-person pronouns are less frequently encountered in online reviews than first-person and second-person pronouns (e.g., Liang et al. 2014). Another reason is that our theoretical framework is developed based on the close association of first-person pronouns with self-focus, and close association of second-person pronouns with other-focus. However, the use of third-person pronouns has no clear association with either self-focus or other-focus, so our theoretical reasoning is not applicable to third-person pronouns. This being said, the role of third-person pronouns

and their interaction with first-person and second-person pronouns are interesting and worthy of future exploration.

Third, we examined the role of personal pronouns in online reviews that involve written words. Online reviews in a written format differ from oral communication in that spoken words carry more cues such as pitch, tempo, loudness, and voice quality (Burgoon et al. 2016). These richer cues might be more able to show empathic concern to the audience than written communication, and similarly more likely to activate the audience's perception of the source's persuasion motives. Thus, our focus on the written language represents a more conservative test of the role of personal pronouns. While it is reasonable to expect second-person (vs. first-person) pronouns to play a greater role in oral communication (e.g., resulting in greater perceived empathic concern), more research is needed to extend the study of personal pronouns to oral communication.

Fourth, we theorized the role of personal pronouns based on its close association with attentional focus. However, we believe the concept of attentional focus itself is very important, and that a difference in attentional focus may be manifested in ways other than personal pronouns. Although attentional focus is not the emphasis of this work, its role in online reviews is an interesting avenue for future research. In addition, how to influence reviewers to be more other-focused is another worthy direction. For example, how to motivate reviewers to be more other-focused when writing a review (e.g., giving them a review template, or implicitly priming them to take the perspective from a prospective consumer)? What kind of reviewers are more likely to be influenced to write an other-focused review? Answers to these questions can better help firms put our findings into action.

Fifth, we explored and ruled out a few additional mechanisms (other than perceived empathic concern and perceived persuasion motives) underlying the main effect of personal pronouns and the moderating effect of review two-sidedness in the supplementary study. However, our results revealed that the effect of personal pronouns may also be explained by readers' perception of reviewer altruism. Although this finding does not conflict with our empathy-based arguments (because altruism is very likely to be a natural consequence of empathic concern), future research should explore the relationship between empathy and related concepts such as altruism and whether they are always in line with each other.

Sixth, our dependent variable in this study is review helpfulness, defined as the extent to which an online review is perceived by consumers to facilitate their purchase decision process. Helpful reviews can provide diagnostic value for consumers to make judgment and decisions. On the other hand, we did not explore more downstream consequences of personal pronouns beyond review helpfulness, such as consumers' attitude and their intention to purchase a product, download an app, or adopt a service. Future research is needed to explore the impact of personal pronouns on these other relevant outcomes during consumers' decision-making process.

Finally, we tested the interpersonal effect of personal pronouns in the context of digital products' reviews, as the reviews that we collected or manipulated across the five studies are describing mobile applications. Although our theoretical framework and arguments based on empathic concern and persuasive motives should apply to persuasive writing in general, future work is needed to examine the external validity of our findings in other online settings such as other types of online reviews (e.g., service reviews and

physical product reviews) and other forms of persuasive writing (e.g., persuasive appeals in advertising).

2.9   Conclusion

Recognizing the critical role of function words in online reviews, we examine how and when a reviewer's use of personal pronouns can influence consumers' helpfulness evaluation of the online review. We propose that the use of second-person (vs. first-person) pronouns does not always increase review helpfulness, and that its overall effect depends on review two-sidedness. Through an archival analysis and four experiments, we provide converging evidence for our theoretical framework. These findings emphasize the importance of studying the role of function words such as personal pronouns in online reviews and open up exciting opportunities for future research.

CHAPTER 3.   CONFIRMATORY OR DISCONFIRMATORY REVIEWS? EXAMINING CONSUMERS' SELECTIVE EXPOSURE IN SEEKING AND EVALUATING ONLINE REVIEWS

3.1   Introduction

Online reviews are playing an increasingly important role in consumers' decision-making process. However, because of the exploding number of reviews that can cause information overload for consumers (Jones et al. 2004), identifying helpful reviews and understanding their determinants have attracted tremendous attention from researchers. Specifically, prior studies have examined the roles of review ratings (e.g., Korfiatis et al. 2012; Sen and Lerman 2007), content characteristics (e.g., Cao et al. 2011; Korfiatis et al. 2008), expressed emotions (e.g., Yin et al. 2014; Yin et al. 2017), reviewer characteristics (e.g., Forman et al. 2008), product type (e.g., Mudambi and Schuff 2010), etc.

While research in this area generally assumed that different consumers would evaluate the same review similarly, recent studies have started to scrutinize this assumption and explore the role of consumers' initial beliefs (Cheung et al. 2009; Qiu et al. 2012; Yin et al. 2016). Notably, Yin et al. (2016) found that consumers' pre-existing beliefs about a product influence their helpfulness evaluation of the product's reviews, and that consumers demonstrate a tendency of confirmation bias – evaluating confirmatory reviews that are consistent with their initial beliefs more favorably than disconfirmatory reviews. Despite their best efforts in addressing endogeneity and

selection bias issues, the nature of their observational data could not provide the definite, causal evidence for the presence of confirmation bias.

In addition to shaping consumers' evaluation of review helpfulness, their initial beliefs may also affect the way they look for reviews. Before consumers read any particular reviews, they need to first be selective in navigating the vast sea of information (see Mathieson and Wall 1982; Woodside and MacDonald 1994). The first impressions consumers form about a product may guide the types of reviews they seek out in a systematic manner. A deeper understanding of how consumers seek information can help review platforms incorporate the demand factor into the ranking calculation of reviews that typically relies solely on readers' helpfulness evaluation, bringing the content most sought after to the forefront (e.g., pushing reviews in higher demand to a more prominent position even if they are relatively new without a helpful score or rated less helpful). Such knowledge can also help product manufacturers to adjust their priority in dealing with reviews based on their likely exposure to prospective consumers.

More importantly, explorations into consumers' information seeking tendencies may uncover possible reasons behind a negativity bias in the impact of online reviews on product sales. Prior research has repeatedly demonstrated that negative reviews hurt product sales to a greater extent than positive reviews help sales (e.g., Basuroy et al. 2003; Cao et al. 2011; Chevalier and Mayzlin 2006). A commonly assumed explanation for this negativity bias is that negative reviews are weighted more heavily and perceived more helpful by consumers. However, recent empirical evidence suggests the opposite: Yin et al. (2016) found that consumers' evaluation of review helpfulness is driven by their initial beliefs about a product rather than review valence *per se*; because the average

rating of most products is positive, overall consumers demonstrate a positivity bias –

evaluating positive reviews (that confirm their positive initial beliefs) more favorably

than negative reviews. How could negative reviews exert a greater impact than positive

reviews on product sales, while the former are generally considered less helpful than the

latter? A very likely reason is the way consumers look for reviews: if consumers prefer to

seek out negative reviews than positive reviews, then negative reviews would get more

exposure and be read by more consumers; if the greater exposure of negative reviews

outweighs their lower helpfulness, then negative reviews could have a greater impact on

product sales than positive reviews even though the former are considered less helpful. In

contrast, if the confirmation bias (and a resulting, overall positivity bias) observed when

consumers evaluate the helpfulness of review information (Yin et al. 2016) is also present

when they seek out reviews to read, then positive rather than negative reviews should

have a greater impact on product sales.

Despite the importance of information seeking in the context of online reviews,

no research to our knowledge has focused on this earlier stage of consumers' decision-

making process or explored the role of consumers' initial beliefs in their information

seeking tendencies. In this paper, we examine the role of consumers' initial beliefs and

their tendencies to prefer confirmatory or disconfirmatory information in both review

seeking and review evaluation stages. Building on and extending the motivated reasoning

literature, we propose that in the information seeking stage, consumers with positive

initial beliefs tend to look for disconfirmatory reviews to read, while consumers with

negative initial beliefs tend to seek out confirmatory reviews; overall, they should

demonstrate a negativity bias in their information seeking tendencies. In the second stage

when consumers read and evaluate particular reviews, however, their motivation to avoid conflict and defend their initial beliefs may propel them to evaluate confirmatory information more favorably regardless of the valence of their initial beliefs. To test our proposed theoretical framework, we conducted three laboratory experiments.

Our paper makes a number of unique contributions to the online reviews literature. First, although we know a great deal about factors driving consumers' evaluation of review helpfulness, this research is the first, to our best knowledge, to examine how consumers seek out reviews to read in an earlier stage of their decision-making process. Second, we propose and find consistent evidence that consumers' tendency to seek out confirmatory or disconfirmatory reviews is dependent on the valence of their initial beliefs, and they demonstrate an overall negativity bias in the information seeking stage. These findings expand our understanding of the roles of consumers' initial beliefs to more stages of their decision-making process, and also suggest the existence of multiple motivations that might drive distinct biases in different stages. Our demonstration of an overall negativity bias when consumers look for reviews to read also provides a plausible explanation for the negativity bias that has been reliably shown in the online word-of-mouth literature to influence product sales. Third, our paper complements prior research demonstrating the existence of confirmation bias in review evaluation (Yin et al. 2016) by providing experimental (and thus causal) evidence for the occurrence of this bias. Our findings also offer important practical implications for product manufacturers and review platforms.

3.2 Literature Review and Hypotheses Development

3.2.1 Consumers' Initial Beliefs

Before consumers seek out and read any reviews of a product, they may have already formed initial beliefs about the product. Because of the critical role that online reviews play in consumers' purchase decisions, most review sites display summary statistics of a product's ratings prominently, including the average rating and number of reviews of the product. These rating profiles can facilitate the formation of consumers' initial beliefs about the product even before they get to any reviews (Yin et al. 2016).

Once formed, consumers' initial beliefs can have significant implications for their subsequent judgment and intentions. Prior studies found that consumers' initial beliefs could impact their evaluation of product reviews (e.g., Cheung et al. 2009). Recent evidence further suggests that consumers have a general preference for confirmatory reviews during information evaluation and evaluate such reviews more favorably (Yin et al. 2016). However, the influence of consumers' initial beliefs may not be limited to the evaluation process, as consumers may be similarly selective when they decide which kinds of reviews to read first. Consumers' preference for confirmatory or disconfirmatory information is labeled selective exposure in the social cognition literature, which we turn to next.

3.2.2 Selective Exposure, Confirmation Bias, and Disconfirmation Bias

Selective exposure refers to individuals' systematic preference for attitude-congruent or attitude-incongruent information (Knobloch-Westerwick et al. 2013).

Substantial evidence from experimental studies showed that people have a general tendency to prefer information that is consistent with their initial beliefs to information that is inconsistent, and this particular tendency is often termed confirmation bias (Jonas et al. 2001; Lazarsfeld et al. 1944). For instance, in a presidential election, voters reveal preferences for political messages that are in line with their political views and leanings (e.g., Chaffee et al. 2001; Stroud 2008). In health communication, people actively avoid messages challenging their beliefs (Case et al. 2005; Knobloch-Westerwick et al. 2013; Pease et al. 2006). In interpersonal relationships, people also tend to seek information that is consistent with their initial beliefs about a target individual (e.g., Snyder 1981; Snyder 1984).

However, the evidence for confirmation bias is not universal, with a number of studies revealing a disconfirmation bias – a preference for attitude-inconsistent information compared to attitude-consistent information (Edwards and Smith 1996; Taber and Lodge 2006). For example, when participants were given synopses of criminal trials and asked to read the defense or the prosecution summation, they preferred to seek out information that contradicts their own opinions (Sears 1965). There was also evidence suggesting that confirmation bias could be attenuated or even reversed when the inconsistent information has higher informational utility (Hastall 2009; Knobloch et al. 2003; Knobloch-Westerwick et al. 2005). Taken together, although confirmation bias is a largely ubiquitous phenomenon, evidence from prior research is not conclusive. In the following, we introduce a prominent account of confirmation and disconfirmation biases based on people's fundamental motivations.

3.2.3 *Accuracy and Defense Motivations*

The motivated reasoning literature from social psychology provides a theoretical framework for explaining the divergent findings with regard to selective exposure (Eagly et al. 1999; Johnson 1994; Prislin and Wood 2005). A basic premise of motivated reasoning is that people's motivations can affect the process of their reasoning – forming beliefs, evaluating evidence, and making decisions (Erdelyi 1974; Festinger 1957). Two fundamental motivations proposed in this literature are accuracy motivation and defense motivation (Kunda 1990). Accuracy motivation refers to one's desire to uncover the truth and form accurate evaluations of stimuli (Chaiken et al. 1989; Kunda 1990), while defense motivation refers to one's desire to defend prior beliefs, attitudes, and behaviors.

Accuracy and defense motivations have been found to influence how people process attitude-inconsistent and attitude-consistent information (e.g., Chaiken et al. 1996; Prislin and Wood 2005; Wyer and Albarracín 2005). On one hand, people tend to show disconfirmation bias when they are motivated to uncover the truth and make good decisions (Chaiken et al. 1989; Hart et al. 2009). Accuracy motivation has been found to reorient people's attention to information utility (Hart et al. 2009), defined as the degree to which the information can be used to make successful decisions (Fischer et al. 2011). Specifically, the accuracy motivation of consumers may drive them to engage with information with greater utility, because such information can better fulfill their ultimate goal of making a good purchase decision (e.g., Fischer and Greitemeyer 2010; Knobloch-Westerwick and Kleinman 2012). Information that is inconsistent with people's existing beliefs should be perceived to have higher utility, because inconsistent information provides more evidence or opinions beyond their knowledge and thus has more

informational value. Compared with attitude-consistent information, inconsistent information is also more salient and more likely to evoke their attention and interest (e.g., Berlyne 1970; David 1996). Therefore, in order to be accurate, consumers should be more likely to prefer disconfirmatory information than confirmatory information.

On the other hand, people reveal confirmation bias when they are motivated to defend their prior beliefs (Chaiken et al. 1989; Hart et al. 2009). Disconfirmatory information that is incompatible with consumers' prior beliefs provokes the negative arousal state of cognitive dissonance (Festinger 1957), arising from the discomfort caused by cognitive conflicts (Beauvois and Joule 1996; Harmon-Jones 2000). Because people generally dislike cognitive dissonance and its associated discomfort, disconfirmatory information is more likely to be refuted and disregarded (Wyer and Frey 1983), or subject to more extensive and critical scrutiny, than confirmatory information (e.g., Ditto and Lopez 1992; Koehler 1993; Kunda 1990). As a result, when experiencing or anticipating cognitive dissonance, people tend to prefer attitude-consistent (compared with attitude-inconsistent) information by assigning more weights to them (Fischer et al. 2011; Frey 1986; Hart et al. 2009). Thus, under defense motivation, consumers are more likely to favor confirmatory information compared with disconfirmatory information. Next, we posit that consumers' accuracy and defense motivations depend on the stage of their decision-making process and, under certain circumstances, the valence of their initial beliefs. We first introduce the two stages of consumers' decision-making process and then explain consumers' distinct motivations in each stage.

3.2.4 Two Stages of Consumers' Decision-Making Process

When consumers are deciding whether to purchase certain products, their decision-making process involves two stages: information seeking and information evaluation. During the information seeking stage, consumers actively search for related information. During the information evaluation stage, consumers evaluate and appraise available information to reform their beliefs and attitudes, which will impact their final decisions (Fischer et al. 2008a). According to several prominent models of consumers' decision-making process (e.g., Mathieson and Wall 1982; Woodside and MacDonald 1994), the information seeking stage is normally followed by the information evaluation stage before consumers make a final decision.

Among prior studies examining selective exposure to information in diverse contexts, most focused on either information seeking or information evaluation. For example, researchers have investigated how individual differences shape the way people seek out information about relationships (e.g., Brannon et al. 2007; Holton and Pyszczynski 1989; Rholes et al. 2007; Sargent 2007). There are also studies exploring selective exposure in the information evaluation stage, such as different decision criteria people use to assess confirmatory and disconfirmatory information (e.g., Carlson and Russo 2001; Ditto and Lopez 1992; Greitemeyer and Schulz-Hardt 2003b; Russo et al. 1998).

However, very few studies have examined selective exposure in both stages simultaneously. In addition, as mentioned earlier, evidence for consumers' selective exposure to information during their decision-making process is not conclusive. Next,

building on the motivated reasoning literature and the unique context of online reviews, we propose that both disconfirmation and confirmation biases are likely to occur during consumers' decision-making process.

3.2.5   Information Seeking Stage

Given the abundance of product options and available information for any purchase decision, consumers often limit their attention and evaluation to a subset of available options (named "consideration set") to simplify their decisions (Roberts and Lattin 1991; Wright and Barbour 1977). Because consumers typically engage in-depth information processing and make final decisions (which product to purchase) among product options that fall into their consideration sets, the determinants of consumers' consideration sets play a fundamental role in their judgment and choice (Shocker et al. 1991). The likelihood of a product option to be included in consumers' consideration sets is determined by a largely rational, cost-benefit analysis; a product is more likely to be included if the perceived benefit of evaluating it exceeds the perceived cost (Roberts and Lattin 1991). Assuming that costs of evaluating all products are the same, consumers should be more likely to include the product options that they have more positive (or less negative) beliefs toward (i.e., expecting these options to have greater utility and bring more potential benefit to them) in their consideration sets.

In our context, consumers can readily form initial beliefs toward product options well before they get to any particular consumer reviews. Specifically, aggregated rating profiles of a product, such as the average and number of ratings, are often prominently displayed along with product options, and they have been found to help consumers form

initial beliefs about the product (Yin et al. 2016). For instance, a product's average rating is perceived by consumers to reflect its quality (De Langhe et al. 2015), while the number of ratings reflects the product's popularity (Chevalier and Mayzlin 2006; Duan et al. 2008). Thus, these salient cues can help consumers form a positive or negative initial belief about the product, which in turn facilitates their decision on whether to include the product in their consideration set.

Because of the close association of the valence of consumers' initial beliefs toward a product option and the likely inclusion of the option in their consideration set, we argue that consumers' initial beliefs also drive distinct motivations. First, consumers with positive initial beliefs are more likely to be motivated by accuracy when seeking out reviews. When consumers form positive initial beliefs about a product based on the positive average rating or a large number of its ratings, they are more likely to place the product in their consideration set and thus purchase the product in the end. At the same time, such initial beliefs developed on the basis of aggregated rating cues are typically not strong or validated, and they cannot help consumers make a choice among similarly rated product options. To avoid making a poor decision under uncertainty, people tend to be more vigilant with a cautious mindset; compelling evidence also suggests that people are motivated more by accuracy in uncertain and ambiguous circumstances (Fischer et al. 2011; Fischer et al. 2008b; Greitemeyer and Schulz-Hardt 2003a). Thus, when consumers seek reviews to primarily reduce the uncertainty about a product that they are very likely to purchase (Dellarocas 2003), they tend to be motivated more by accuracy in order to increase the chance of a wiser and better decision.

Combining this with earlier arguments that accuracy motivation prompts consumers to seek out disconfirmatory reviews, we hypothesize that in the information seeking stage, disconfirmation bias is likely to occur when consumers form positive initial beliefs about a product. Specifically, after consumers form a positive initial belief towards a product based on its rating profiles, they are more likely to be motivated by accuracy to seek out information with greater utility and informational value. Thus, their accuracy motivation should drive them to search more for disconfirmatory reviews than confirmatory reviews. Taken together, we propose the following hypothesis.

Hypothesis 1a: Consumers with positive initial beliefs about a product would prefer to read disconfirmatory (negative) reviews than confirmatory (positive) reviews.

Next, we posit that consumers with negative initial beliefs are more likely to be motivated by defense when seeking out reviews. Consumers can form a negative initial belief about a product based on the product's rating profiles, such as when the product has a lower or negative average rating (Forman et al. 2008). With such a negative impression about a product, consumers are less likely to put the product in their consideration set (Shocker et al. 1991). When consumers exclude a product from their consideration set, the likelihood of purchasing the product is fairly low because final purchase decisions are typically made among the options in the consideration set. The largely strong and certain nature of consumers' negative initial beliefs should activate their defense motivation that typically accompanies strong beliefs and attitudes (Brechan 2002). In addition, because consumers with negative initial beliefs about a product are less likely to incorporate the product in the consideration set, they would have less vested

74

interest in uncovering the true quality of the product and thus have lower motivation to be accurate.

Combining this with earlier arguments that defense motivation leads consumers to prefer confirmatory information, we hypothesize that in information seeking, confirmation bias is more likely when consumers form negative initial beliefs about a product. After consumers form a negative initial belief towards a product, their motivation to defend their existing impression about the product should overweigh their motivation to uncover the truth of product quality. Thus, the predominance of their defense motivation should drive them to search more for confirmatory reviews than disconfirmatory reviews. In all, we propose the following hypothesis.

Hypothesis 1b: Consumers with negative initial beliefs about a product would prefer to read confirmatory (negative) reviews than disconfirmatory (positive) reviews.

It is worth pointing out that our hypotheses of disconfirmation and confirmation biases in the information seeking stage are not in conflict with a general tendency of people to seek negative information (termed "negativity bias") (Rozin and Royzman 2001). We posit that consumers with positive initial beliefs would prefer to read disconfirmatory (negative) reviews, while consumers who have negative initial beliefs would prefer to read confirmatory (negative) reviews. Thus, our accounts based on differential motivations provide a plausible explanation for the presence of negativity bias when consumers look for reviews to read.

3.2.6 Information Evaluation Stage

After consumers have a chance to read the actual review content, we posit that confirmation bias is likely to occur regardless of the valence (positive or negative) of their initial beliefs because consumers' defense motivation prevails when they read and evaluate the review content. In this stage, consumers are exposed to the actual content of individual reviews. When they encounter information that is incongruent with and directly contradicting their initial beliefs and attitudes, such conflict can cause discomfort, a form of psychological stress that people generally dislike (Festinger 1957). The presence of such discomfort can trigger consumers' motivation to reduce it and defend their existing opinions (e.g., Beauvois and Joule 1996; Harmon-Jones 2000). The heightened likelihood of encountering actual conflict in information evaluation (as opposed to information seeking in which no conflict was experienced) should contribute to the dominance of consumers' defense motivation during this stage.

Integrating these arguments, we predict that confirmation bias is likely to occur when consumers evaluate the helpfulness of reviews no matter whether their initial beliefs are positive or negative. Consumers are more likely to encounter cognitive conflict and discomfort because of their direct access to substantive content of particular reviews, and activate their defense motivation as a result. Fueled by this motivation, consumers may refute or discount disconfirmatory information and evaluate confirmatory reviews more favorably. As such, we propose the following hypothesis.

Hypothesis 2: Confirmatory reviews are perceived to be more helpful than disconfirmatory reviews.

76

To test these hypotheses, we conducted three controlled experiments. In the first two studies, we examined H1a and H2 when the valence of consumers' initial beliefs was positive. Study 1 was a hypothetical online decision-making task, in which participants formed positive initial beliefs about a product before they were presented with a number of the product' reviews to choose and subsequently read. Study 2 utilized a more realistic scenario and replicated the first study's findings. In the final study, we manipulated the valence of participants' initial beliefs about the product to test the full set of our hypotheses.

3.3 Study 1

In Study 1, we designed an experiment in which participants formed positive initial beliefs about a product before selecting and reading its reviews. Specifically, subjects were presented with the rating profiles of two wireless mouse products, and then they were asked to pick one that they are more likely to purchase. We varied average ratings of the two product options so that one option would appear superior to the other and subjects would develop a positive impression about the superior option. After participants (presumably) picked the superior product for further investigation, they were asked to select 3 out of 6 reviews (3 positive and 3 negative) of this product to read based on the reviews' titles, read the content of selected reviews and then report their helpfulness evaluation of the reviews. We captured participants' selective exposure in information seeking by comparing the number of selected reviews that confirm and disconfirm their positive initial belief toward the superior product. Moreover, we measured participants' selective exposure in information evaluation by comparing their helpfulness evaluations of the selected confirmatory and disconfirmatory reviews.

77

3.3.1   Stimulus Materials

In this experiment, we selected the compact and foldable wireless mouse because it is familiar and useful to the undergraduate participants. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. A compact and foldable wireless mouse allows people to easily take it with them anywhere they go.

We developed stimuli for this experiment in two steps. In the first step, we prepared 6 review titles that differ in valence but not in extremity. We began with 12 reviews titles (6 positive and 6 negative) after consulting actual review titles of similar products from Amazon.com. To identify positive and negative review titles that are equally extreme, we conducted a pretest (see Appendix C for detailed instructions) and recruited 36 subjects from Amazon Mechanical Turk (MTurk). Each pretest subject was asked to read the 12 review titles, one at a time, and rate the extremity of each title along a 9-point scale (ranging from "not at all negative/positive" to "very negative/positive") adapted from Lee et al. (2009). Based on the results of paired-sample t-tests, we selected 3 positive review titles ("Attractive," "Terrific," and "Wise choice") and 3 negative ones ("It's worthless," "Depressing purchase," and "Disturbing"); comparisons in all pairs of positive versus negative titles yielded a t-value of at most 1.650 with a p-value of at least .108. Therefore, the 6 review titles used in this experiment are not significantly different in their extremity.

In the second step, we prepared 3 sets of text reviews (with a positive version and a negative version in each set) so that the two versions within each set are equivalent in extremity, and different review sets are equivalent in terms of information quantity,

quality, reading difficulty, appropriateness, and realism. We started with 6 sets of text reviews by again consulting real reviews of similar products from Amazon.com. Within each set, we first created a positive review, and then constructed a corresponding negative review by changing its valence (e.g., using antonyms and adding negations) while holding the substantial content identical. We also kept constant the number of words between the two versions in each review set to reduce the likelihood of possible confounds; the only difference between the two versions is valence. Then we conducted another pretest, recruited 72 subjects from MTurk, and asked them to read and evaluate 6 reviews chosen from different sets, one review at a time. Each subject was randomly assigned to read one version (positive or negative) of the reviews in each set. After reading each review, subjects were asked to report their evaluation of its 1) extremity using the same item as in the pretest of review titles, 2) information quantity using two items adapted from Gao et al. (2012), 3) quality using three items adapted from McKinney et al. (2002), 4) reading difficulty using two items adapted from Hall and Hanna (2004) and Ermakova et al. (2014), 5) appropriateness using two items adapted from Glikson et al. (2017), and 6) realism using two items adapted from Mafael et al. (2016). All items were presented along 9-point scales (see Appendix D for all the measures). Based on the results of independent-samples t-tests of extremity and paired-samples t-tests of all other variables (e.g., information quantity), we selected 3 sets of reviews that satisfy our criteria (see Table 3-1): comparisons in extremity of two review versions within each set yielded a t-value of at most 1.380 with a p-value of at least .172; comparisons in all other relevant variables across different sets of reviews yielded a t-value of at most 1.587 with a p-value of at least .117. Therefore, the 3 sets of chosen

reviews used in this experiment are not significantly different either in their extremity

between positive and negative versions of the same review set, or in the other relevant

aspects (e.g., information quantity, quality, etc.) across review sets.

Table 3-1 Content of Reviews in the 3 Sets.

Set #	Positive Version	Negative Version
1	This is a great mouse and it works well. The mouse has the curved left side for the thumb, so it's very comfortable. Moreover, it allows me to change how quickly the cursor moves across my screen.	This is a worthless mouse and it doesn't work well. The mouse doesn't have the curved left side for the thumb, so it's very uncomfortable. Moreover, it doesn't allow me to change how quickly the cursor moves across my screen.
2	The mouse functions well. One feature that I found useful for saving battery life is the mouse turns off automatically after a long time of non-use. It is convenient for someone who walks away from their computer often.	The mouse functions poorly. One feature that I found harmful for saving battery life is the mouse doesn't turn off automatically after a long time of non-use. It isn't convenient for someone who walks away from their computer often.
3	Good value for the price. It includes a battery with the product, so you can use it immediately. It connects to my laptop very quickly. And it is responsive without any lag when I move it.	Poor value for the price. It doesn't include a battery with the product, so you cannot use it immediately. It connects to my laptop very slowly. And it isn't responsive with lags when I move it.

3.3.2 Procedure

36 undergraduate students from a Midwest U.S. university participated in this

experiment in exchange for extra credit. In the cover story, participants were asked to

imagine that they were planning to purchase a compact and foldable wireless mouse from

Amazon.com, and their search returned two different wireless mice with the same price

of $23.99. Then they were asked to read the rating profiles of the two options. The two

product options had both accumulated hundreds of reviews, but their average ratings

were 2 and 4 stars, respectively, with the latter being the superior option. To mitigate

location effects, we counterbalanced whether the superior option appeared on the left or

right of the screen. An example of rating profiles is presented in Figure 3-1.

Mouse A	Mouse B
Price: $23.99	Price: $23.99
Average Rating: ★★☆☆☆	Average Rating: ★★★★☆
Based on **hundreds of reviews** from prior customers	Based on **hundreds of reviews** from prior customers

Figure 3-1 Rating Profile of Two Mouse Options.

After observing the rating profiles of two options side by side, participants were

asked about their initial belief towards each product option to facilitate the formation of

their initial impressions on the products. Afterwards, they were told that they were in a

hurry and only had time to read reviews from one of the two product options. Thus, they

were asked to choose one from the two product options along an 8-point scale (ranging

from "definitely choose Mouse A" to "definitely choose Mouse B"). 35 out of 36

participants preferred the wireless mouse with the 4-star average rating. We retained only

these 35 participants in our main analyses to assure that they had developed positive

initial beliefs about the selected wireless mouse before they got exposed to the review

titles and reviews.

Next, in the information seeking stage, participants were asked to select reviews

that they preferred to read. They were shown the titles of 6 most recent reviews of their

81

selected product option. The participants were told that they did not have enough time to read all the reviews, and that they needed to choose the 3 reviews that they were most interested in reading based on the review titles. The 6 review titles differ in valence (3 positive and 3 negative) but not in extremity based on our pretest results. The order of the 6 review titles was randomized.

Finally, in the information evaluation stage, participants read the 3 text reviews that correspond to the titles they selected in the previous stage and reported their evaluations of each review. These 3 text reviews were selected from the 3 sets of text reviews we pretested earlier, one version from each set. The valence version (positive or negative) in each review set was determined by the valence of the selected review titles. For example, if participants chose 2 negative and 1 positive review titles, they would see 2 negative and 1 positive text reviews, one from each review set. To strengthen the valence manipulation, we also displayed the review rating (5 stars for the positive review and 1 star for the negative review) and review title to go along with each text review. An example of 3 reviews is illustrated in Figure 3-2. Participants were then asked to report their perception of helpfulness for each review. Perceived review helpfulness was measured using a 9-point scale with two items adapted from Sen and Lerman (2007) and Chen and Lurie (2013): "Assuming that you were thinking about purchasing Mouse A/B in real life, how likely would you be to use this review in your decision-making?" (ranging from "very unlikely" to "very likely") and "How much influence would this review have on your decision?" (ranging from "very little influence" to "a great deal of influence"). See Appendix E for detailed instructions and measures.

Figure 3-2 Review Stimuli.

3.3.3   Results

First, we investigated the direction of selective exposure when consumers with positive initial beliefs seek more information. Because their initial belief toward the superior product option should be positive, positive review titles represent confirmatory information (which is consistent with participants' initial belief), and negative review titles represent disconfirmatory information. There are 4 possibilities with regard to the number of positive and negative review titles that a participant could select – 3 positive titles, 2 positive and 1 negative titles, 1 positive and 2 negative titles, and 3 negative titles. We measured the direction of selective exposure in information seeking by

83

comparing the number of selected positive review titles with the number of selected negative review titles. A repeated-measures ANOVA analysis showed that participants preferred to read negative reviews rather than positive reviews ($M = 1.83$ vs. 1.17, $F(1, 34) = 7.570$, $p = .009$), providing evidence for disconfirmation bias in the information seeking stage as hypothesized in H1a (see the bar chart in Figure 3-3).

Next, we examined the direction of selective exposure when consumers evaluate the helpfulness of reviews after reading their content. For this analysis, we retained 29 (out of 35) participants who read both positive and negative reviews (i.e., 2 positive and 1 negative reviews, or 1 positive and 2 negative reviews) because a within-subject comparison is only plausible in such cases. A repeated-measures ANOVA analysis showed that positive reviews were perceived to be more helpful than negative reviews ($M = 7.28$ vs. 5.69, $F(1, 28) = 17.004$, $p < .001$), providing evidence for confirmation bias in the information evaluation stage as hypothesized in H2 (see the solid line in Figure 3-3).
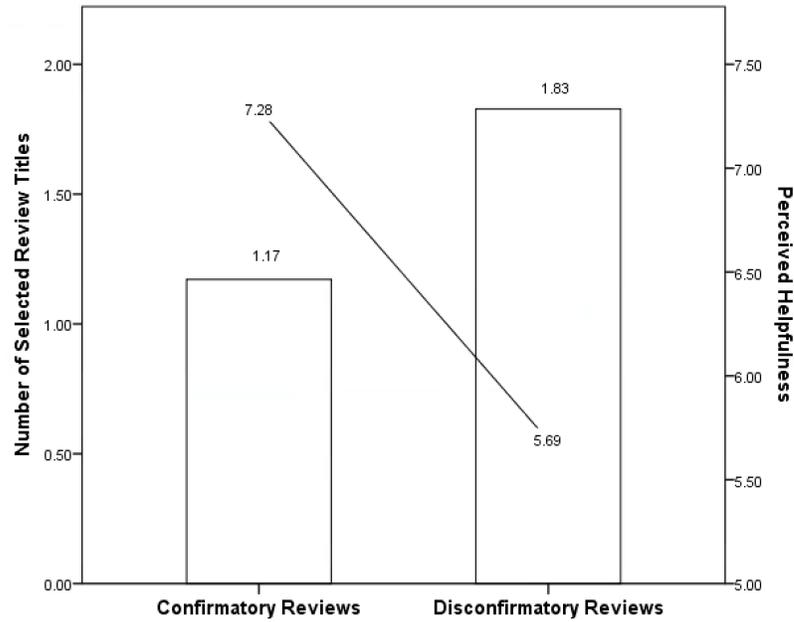
Figure 3-3 Results of Confirmation and Disconfirmation Biases in Study 1.

3.3.4 Discussion

In this study, we conducted an experiment to test the direction of selective exposure in the two stages of information seeking and evaluation when consumers have positive initial beliefs about a product. We found evidence suggesting that consumers seek and prefer to read disconfirmatory reviews that go against their pre-existing positive impression of a product, but after reading the review content, they evaluate confirmatory reviews more favorably. These results provide initial evidences for H1a and H2.

One notable limitation of this study is its artificiality: participants were asked to select from two product options with one being clearly superior to the other (4-star vs. 2-star on average). Because it is not common for a product to have a 2-star average rating, participants would almost certainly purchase the option with the 4-star average rating (if

no other options are available) and develop very strong positive belief toward this superior option. It will be interesting to explore whether differential selective exposure observed in this study could be replicated in a more realistic scenario where the positive initial beliefs of participants was manipulated more subtly.

3.4 Study 2

The primary purpose of Study 2 was to replicate the main findings of Study 1 in a more realistic scenario. This study followed the similar procedure as in Study 1, except that we kept the average rating identical between the two product options, but varied the number of reviews. Multiple products with similar average ratings but differing number of reviews are more likely to happen in real-life shopping scenarios.

3.4.1 Procedure

39 undergraduate students took part in this study for extra credit. The cover story and procedure were similar to those of Study 1, with one major exception. We constructed rating profiles of two product options with the same (4-star) average rating and the same price, but they differed in the number of reviews – one has 15 reviews and the other has 1730 reviews. After observing rating profiles of both options side by side, subjects were asked about their initial beliefs and then asked to select 3 out of 6 review titles of the somewhat superior product option (with 1730 reviews). An example of rating profiles is presented in Figure 3-4. The rest of the procedure was identical to that of Study 1. See Appendix F for detailed instructions and measures.

Mouse A	Mouse B
Price: $23.99	Price: $23.99
Average Rating: ★★★★☆	Average Rating: ★★★★☆
Based on **15 reviews** from prior customers	Based on **1730 reviews** from prior customers

Figure 3-4 Rating Profile Stimuli.

3.4.2 Results

First, we investigated the direction of selective exposure when consumers with positive initial beliefs seek more information of the superior product option. We compared the number of selected positive review titles with the number of selected negative review titles in a repeated-measures ANOVA analysis. Results showed that subjects preferred to read negative reviews rather than positive reviews ($M = 1.85$ vs. 1.15, $F(1, 38) = 9.308$, $p = .004$), providing evidence for H1a (see the bar charts in Figure 3-5).

Next, we explored selective exposure during consumers' information evaluation stage. As in Study 1, we used only 32 (out of 39) subjects who selected both positive and negative review titles. A repeated-measures ANOVA analysis revealed that positive reviews were perceived to be more helpful than negative reviews ($M = 6.58$ vs. 5.13, $F(1, 31) = 7.591$, $p = .010$), providing support for H2 (see the solid line in Figure 3-5).
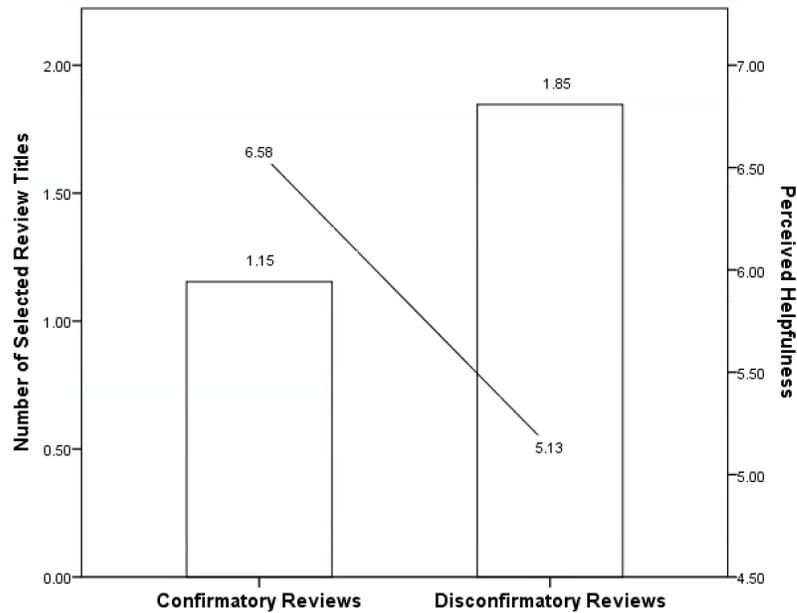
Figure 3-5 Results of Confirmation and Disconfirmation Biases in Study 2.

3.4.3   Discussion

In Study 2, we replicated the findings of the first study by utilizing a more realistic scenario and more subtle manipulation of consumers' positive initial beliefs. In line with H1a and H2, we found consistent evidence that with positive initial beliefs, consumers prefer to read disconfirmatory reviews in the information seeking stage, but they perceive confirmatory reviews to be more helpful in the information evaluation stage.

In both studies, we fixed consumers' initial beliefs toward a product at a positive level, because consumers are more likely to consult reviews of products that are in their consideration set and thus more likely to be purchased. However, such a design precludes us from testing H1b, a theoretically important condition in which consumers with negative initial impressions on a product seek and evaluate reviews of the product.

Therefore, we designed the final study to address this limitation and extend previous findings.

3.5 Study 3

In Study 3, we manipulated consumers' initial beliefs at two levels (positive and negative) and examined their selective exposure during both information seeking and information evaluation stages.

3.5.1 Procedure

103 undergraduate students from a southern U.S. university participated in this study for extra credit. The cover story and procedure were similar to those of Study 1, with one major exception: after observing rating profiles of two product options (2-star and 4-star on average) side by side and answering questions about their initial belief towards each product, they were told that the product on the left appeared first in their search result and caught their attention first, promoting them to check out its reviews first. Because we randomized the location of 2-star vs. 4-star product options, a half of the subjects were assigned to the negative valence condition and asked to seek out and evaluate reviews of the 2-star product while the other half were assigned to the positive valence condition. The rest of the procedure was identical to that of Study 1.

3.5.2 Results

Before further analysis, we conducted a manipulation check to ensure that our manipulation of initial beliefs was successful. The initial belief was measured using a 9-point scale with three items adapted from Darke and Ritchie (2007) (see Appendix G for

the measure). Based on ANOVA analysis, results showed that subjects' initial belief towards the 2-star product was significantly lower than that towards the 4-star product ($M = 2.69$ vs. 7.85, $F(1, 101) = 418.273$, $p < .001$). Thus, our manipulation of the valence of initial beliefs was deemed successful.

First, we examined whether the direction of selective exposure depends on the valence of consumers' initial beliefs toward a product when they seek more information about the product. We conducted a mixed ANOVA analysis, with the nature of selected review titles (confirmatory vs. disconfirmatory) entered as a within-subjects factor, and the valence of initial beliefs as a between-subjects factor. Results revealed that the interaction between consumers' initial belief valence and review title nature was significant ($F(1, 101) = 37.746$, $p < .001$). For the positive initial beliefs condition, pairwise comparisons showed that participants preferred to read disconfirmatory (negative) reviews rather than confirmatory (positive) reviews ($M = 1.86$ vs. 1.14, $t(50) = 3.902$, $p < .001$), providing consistent evidence for disconfirmation bias in the information seeking stage as hypothesized in H1a (see the white bar in Figure 3-6). For the negative initial beliefs condition, pairwise comparisons showed that participants preferred to read confirmatory (negative) reviews rather than disconfirmatory (positive) reviews ($M = 1.98$ vs. 1.02, $t(51) = 4.766$, $p < .001$), providing evidence for confirmation bias in the information seeking stage as hypothesized in H1b (see the black bar in Figure 3-6).[5]

---

[5] As a supplementary analysis, we used the valence of the review title instead of its confirmatory (vs. disconfirmatory) nature as a within-subject factor, and the interaction of this factor and initial belief valence did not reach significance ($F(1, 101) = 0.739$, $p = .392$). Thus, consumers consistently demonstrated a tendency of negativity bias when they seek reviews, although this tendency is in line with our predictions in H1a and H1b.
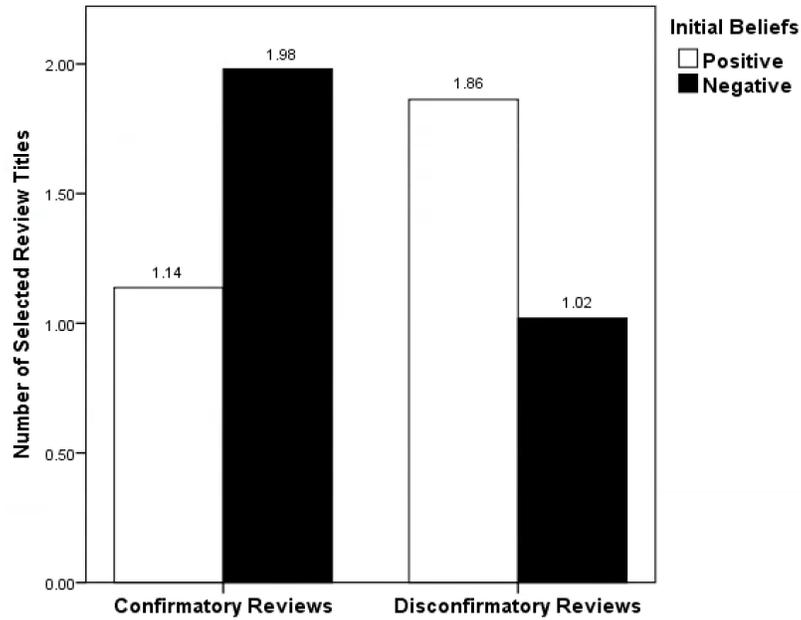
Figure 3-6 Results of Confirmation and Disconfirmation Biases During Information Seeking in Study 3.

Next, we investigated the direction of selective exposure during consumers' information evaluation stage, and whether it depends on the valence of their initial beliefs. As in Study 1, we used only subjects who read both positive and negative reviews (N = 82; 39 out of 51 subjects in positive initial beliefs condition, and 43 out of 52 subjects in negative initial beliefs condition). Results from a mixed ANOVA revealed that the interaction between consumers' initial belief valence and confirmatory (vs. disconfirmatory) nature of reviews did not reach significance ($F(1, 80) = 2.299$, $p = .133$). Moreover, confirmatory reviews were rated as significantly more helpful than disconfirmatory reviews ($M = 6.59$ vs. 5.41, $F(1, 80) = 12.955$, $p = .001$), providing additional evidence for confirmation bias in the information evaluation stage as hypothesized in H2.

3.5.3   Discussion

In Study 3, we replicated the findings of the first two studies and also tested the full set of hypotheses by varying the valence of consumers' initial beliefs. In line with H1a and H1b, we found consistent evidence that consumers with positive initial beliefs prefer to read disconfirmatory reviews, while those with negative initial beliefs prefer to read confirmatory reviews. Interestingly, this pattern of results revealed a consistent tendency of consumers to look for negative reviews (negativity bias) in the information seeking stage. In the information evaluation stage, however, consumers perceive confirmatory reviews to be more helpful than disconfirmatory reviews regardless of the valence of their initial beliefs, supporting H2.

3.6   General Discussion

Drawing on motivated reasoning literature, we hypothesize that in the information seeking stage of consumers' decision-making process, those with positive initial beliefs would be motivated primarily by accuracy and prefer to read disconfirmatory reviews, while those with negative initial beliefs would be motivated by defense and seek out confirmatory reviews instead; overall, they would demonstrate a negativity bias in information seeking as a result. In the information evaluation stage, we propose that the dominance of defense motivation would compel consumers to evaluate confirmatory reviews more favorably regardless of the valence of their initial beliefs. We conducted three experimental studies and provided converging evidence for these hypotheses.

3.6.1 Theoretical Implications

Our paper makes several unique contributions to the online reviews literature. First, while prior research has focused mostly on factors that influence consumers' evaluation of review helpfulness *after* they read the review content, this paper is among the first to explore how consumers seek reviews *before* they actually read any of them. Information seeking is a critical initial step before consumers read and evaluate any particular piece of information (e.g., Fischer et al. 2005; Mathieson and Wall 1982). Despite its importance, little research has examined this earlier stage possibly due to a lack of secondary data about consumers' review seeking tendencies. Our research not only advanced two hypotheses regarding consumers' selective exposure to information in this stage, but also utilized experimental methods and carefully pretested stimuli to capture consumers' information seeking tendencies and test the hypotheses. Thus, our examination of consumers' selective exposure *before* they read any particular reviews extends our understanding of their decision-making process beyond merely review evaluation, and opens up exciting opportunities for future research to examine consumers' information seeking behavior.

Second, this paper provides additional evidence for the important role of consumers' initial beliefs, and introduces two distinct motivations (i.e., accuracy and defense motivations) that can give rise to either confirmation or disconfirmation bias under different situations. Most relevant to our research, Yin et al. (2016) demonstrated empirically that consumers' initial beliefs influence their judgment of review helpfulness and that they evaluate confirmatory reviews more favorably than disconfirmatory reviews (i.e., confirmation bias). Complementing and extending this work, our research reveals

93

that consumers' initial beliefs can also influence their review seeking tendencies. More importantly, we reasoned and showed that confirmation bias found in the review evaluation stage cannot be blindly generalized to the review seeking stage, and that consumers may activate distinct motivations for information searching depending on the valence of their initial beliefs. Specifically, when consumers develop positive initial beliefs toward a product, their predominate accuracy motivation (e.g., seeking the "truth" and being accurate) may drive them to look for disconfirmatory reviews to read (see Gilovich et al. 1993; Knobloch-Westerwick and Kleinman 2012), but when consumers form negative initial beliefs, their dominate defense motivation (e.g., defending existing beliefs and attitudes) may drive them to seek out confirmatory reviews to read (see Chaiken et al. 1989; Hart et al. 2009). In other words, both confirmation and disconfirmation biases can occur as consumers look for more information, and the exact type of biases they engage is guided by the valence of their initial beliefs and the corresponding motivation activated at the time. These findings reveal the importance and value of tapping into consumers' fundamental motivations to understand their behavior of navigating the vast amount of information available from online reviews.

Third, our results provide a possible explanation for the well-established negativity bias in terms of product sales, and help reconcile its contradiction with a recently demonstrated confirmation bias in review helpfulness evaluation. Negative reviews have been found to have greater influence on product sales than positive reviews (e.g., Basuroy et al. 2003; Cao et al. 2011; Chevalier and Mayzlin 2006). A commonly assumed explanation for this negativity bias is that negative reviews are perceived by consumers as more helpful than positive reviews. However, Yin et al. (2016) provided

empirical evidence for a confirmation bias in review evaluation; because the average rating of most products is positive, consumers have been found to evaluate positive reviews (that confirm their positive initial beliefs) more favorably in most cases. Our proposed theoretical framework and findings provide a possible answer to this puzzle. Essentially, consumers engage in both information seeking and information evaluation before they make a purchase decision; thus, the negativity bias universally observed at the product level could be caused by a greater value assigned to negative reviews, or a greater exposure of negative reviews than positive reviews. Our examination revealed a consistent tendency of consumers to search for negative reviews, no matter whether they have formed a positive or negative initial impression of the product. Therefore, negativity bias might arise because negative reviews get more exposure and are consulted by more consumers, not because negative reviews are perceived more diagnostic by consumers.

Finally, our paper provides the first experimental evidence for consumers' confirmation bias when they read and evaluate online reviews. Yin et al. (2016) utilized a panel data set of app reviews from Apple's App Store, and found that consumers demonstrate a tendency to evaluate confirmatory reviews more favorably than disconfirmatory reviews. However, the archival nature of their data set precludes them from making any causal claims about the presence of confirmation bias. Answering their call for "laboratory experiments [that] could be an alternative method to answer these questions", our research not only replicates their findings, but also provides the first causal evidence for the existence of confirmation bias in review evaluation.

3.6.2   Practical Implications

Our findings also offer a number of practical implications for product manufacturers and review platforms. First, when product manufacturers establish their priority and strategies of dealing with the tremendous amount of online reviews (e.g., responding to reviewer comments), they should take into account the number of consumers who are likely to be exposed to a review (and thus influenced by the review) in addition to the review's perceived helpfulness. If a product's average rating is positive, then negative reviews of the product would be discounted as unhelpful because negative information contradicts consumers' initial beliefs formed on the basis of the average rating. As a result, such negative reviews are less likely than positive reviews to get into the list of "most helpful" reviews or be prominently displayed on the product page, and a rational product manufacturer may disregard such reviews and focus their attention and resources on the most helpful ones. However, our findings suggest that this strategy might be misguided, because negative reviews in this case contradict consumers' positive initial beliefs and should get more exposure (i.e., being sought after and read by more consumers; see H1a). In addition, dealing with negative reviews proactively is an unequivocally superior strategy only when a product's average rating is negative, because negative reviews are both sought after (confirming consumers' negative initial beliefs as hypothesized in H1b) and rated more helpful (see H2) in this situation. Therefore, when product manufacturers prioritize their efforts in dealing with distinct types of reviews, they should take a more balanced view, considering both the perceived value of a review and the number of prospective consumers likely to be exposed to the review.

96

Second, review platforms such as Amazon may need to reconsider the effectiveness of highlighting the most helpful reviews, and balance consumers' diverse interests at different stages of their decision-making process. Highlighting the reviews rated by others as helpful might bring more confirmatory reviews to the forefront, as one reason behind such helpful reviews is that they are in line with consumers' initial beliefs formed on the basis of the product's average rating and other rating statistics. However, our findings suggest that the review helpfulness metric might not be the only factor that review platforms should incorporate to highlight and sort product reviews. Instead, negative reviews are what consumers actively look for regardless of the valence of their initial beliefs. Although listing the most helpful reviews by default is an intuitive and efficient strategy for review platforms to implement, this strategy neglects the potential of negative reviews to be sought after by more consumers and exert greater impact on product sales. Note that Amazon does provide one "top positive review" and one "top critical review" after consumers click on "see all verified purchase reviews" at the end of the most helpful reviews on the product page. Nevertheless, our findings suggest that displaying negative reviews more prominently along with most helpful reviews may help consumers the most as they navigate the complex process of making a purchase decision.

3.6.3   *Limitations and Future Research*

Our paper also has a few limitations for future examination. First, we fixed the average rating of the treatment product to be 2 or 4 stars in three studies, because our primary interest is the valance (positive or negative) of consumers' initial beliefs. However, this design precludes us from examining situations where consumers have neutral or mixed initial beliefs about a product, such as when the product's average rating

is 3 stars or if reviewers have very divergent opinions (characterized by high dispersion of ratings). It would be interesting to investigate whether consumers are still selective in seeking and judging positive versus negative information when they have neutral or mixed initial impressions on a product. In addition, consumers' confidence in their initial beliefs may influence the strength of their accuracy versus defense motivations when searching for and evaluating reviews. Future research is needed to answer these intriguing questions.

Second, although our theoretical framework built on the two-stage decision-making process of consumers provides a plausible explanation for the negativity bias at the product level, other possibilities exist that warrant further investigation. For example, one possibility is that helpful reviews may not always be persuasive, and that consumers' attitude toward a product might be "swayed" by particular characteristics of the reviews even when those reviews are deemed unhelpful (see Liu and Karahanna 2017). Given the lack of research exploring the association of review helpfulness with consumer attitude and decision-making, this is a fertile area that is worth pursuing.

Finally, our findings provided evidence for the presence of both confirmation and disconfirmation biases during consumers' decision-making process in the context of product reviews. Although our theoretical framework could apply to the general decision-making process of consumers, future work is necessary to test the external validity of our findings in other contexts such as other types of online reviews (e.g., retailer reviews) and other settings such as political voting.

3.7　Conclusion

In keeping with recent research emphasizing the role of consumers' initial beliefs in their judgment of online reviews, we examine consumers' tendency to prefer confirmatory information or disconfirmation information in both stages of review seeking and review evaluation during the decision-making process. Drawing on the motivated reasoning literature, we propose that in the information seeking stage, consumers with positive initial beliefs are more likely to search for disconfirmatory (negative) reviews to read, while consumers with negative initial beliefs are more likely to seek out confirmatory (negative) reviews; in the information evaluation stage, consumers generally evaluate confirmatory reviews to be more helpful. Through three experiments, we find converging evidence for our hypotheses. These findings highlight the critical role of consumers' initial beliefs and their distinct motivations in the decision-making process.

CHAPTER 4.   ANCHORING OR SWAYING? THE IMPACT OF OVERALL AVERAGE RATING VS. MOST ACCESSIBLE REVIEWS IN ONLINE WORD-OF-MOUTH

4.1   Introduction

Online reviews and ratings play an increasingly important role in consumers' decision-making process. Many online retailers and review platforms allow consumers to share their experiences with and opinions of a product in text reviews and to assign a rating (typically 1 to 5 stars) to the product. A higher rating indicates a more positive evaluation of the product. In order to help prospective consumers easily gauge product quality, retailers and review platforms commonly display the product's overall average rating along with other summarized rating statistics (e.g. the total number of ratings, the distribution of ratings) in more prominent places than individual reviews. Prior research found that aggregated rating statistics and individual reviews can each impact consumers' decision-making (e.g., Chevalier and Mayzlin 2006; Mudambi and Schuff 2010).

An implicit assumption among researchers and practitioners is that a product's overall average rating and other summary rating statistics matter more than individual reviews for predicting product sales. For example, large-scale meta-analyses of empirical evidence on determinants of product sales focused on the role of aggregated rating statistics rather than individual reviews (see Babić Rosario et al. 2016; Floyd et al. 2014). In practice, most online retailers and review sites display these aggregated rating statistics not only at the top of product review pages, but also at more prominent places (such as product listing

pages) where no individual reviews are visible. In particular, the overall average rating is the most important cue for consumers to infer the quality of a product (De Langhe et al. 2015), because it is what consumers see first to form an initial impression of the product (Yin et al. 2016). Further, the overall average rating incorporates historical information from all the posted reviews and is therefore a more comprehensive measure of product quality than individual reviews. This view is also in line with the anchoring effect, such that the average product rating can act as an anchor—a salient, initially presented value—before consumers make necessary adjustments based on individual reviews, and their final judgments would be close to the anchor (see Tversky and Kahneman 1974). Thus, this view predicts that the overall average rating of products is the primary driver of consumers' purchase decisions.

However, individual reviews can also affect consumers' purchase decisions. In addition to seeing aggregated rating statistics, consumers also consult individual reviews during the decision-making process. Prior literature studying individual reviews has examined the influence of rating and review text characteristics on the helpfulness of reviews because separating helpful from unhelpful reviews has both theoretical and practical implications (e.g., Mudambi and Schuff 2010; Yin et al. 2017; Yin et al. 2016). However, very few studies have connected individual reviews to consumers' intention to purchase the product or product sales (as an exception, see Shoham et al. 2017).

In addition, no research to our knowledge has directly compared the impact of aggregated rating statistics (derived from all historical ratings) with that of the most accessible individual reviews (those that appear on the first page of reviews) on product sales. In this paper, we revisit the commonly accepted, implicit assumption that a product's

average rating matters more than individual reviews for product sales, and we explore whether and when consumers rely more on individual reviews in decision making. If individual reviews can easily sway product sales, then the current practice of review platforms, product manufacturers and retailers of highlighting the product's overall average rating and other summary rating statistics might not be impactful.

Drawing on the heuristics-and-biases literature (Tversky and Kahneman 1974), we propose two competing hypotheses on the relative impacts of a product's overall average rating and its most accessible reviews on consumers' purchase intention and product sales. On one hand, the *overall* average rating may matter more than individual reviews (namely the anchoring effect) because the product's average rating can serve as a salient anchor and bias subsequent consumer decision making toward the anchor. On the other hand, the average rating of the most accessible reviews may better predict product sales (namely the swaying effect) because prominently displayed individual reviews are more cognitively accessible (and thus more influential) for consumers. In addition, building on the metacognition literature (Smith and Swinyard 1988), we posit that a product's rating dispersion that shapes consumers' confidence in their initial beliefs about the product (Yin et al. 2016) can moderate the likelihood of anchoring and swaying effects. Utilizing two carefully designed experiments and an archival panel dataset collected daily from Apple's App Store over a two-month period, we find evidence supporting an overall swaying (versus anchoring) effect and the moderating role of products' rating dispersion.

Our study has a number of unique theoretical contributions. First, we challenge the widely accepted assumption that a product's overall average rating is the primary source of input in consumer decision making. Among the first attempts to compare the effects of

overall average rating and easily accessible individual reviews, this paper provides both experimental Online reviews and ratings play an increasingly important role in consumers' decision-making process. Many online retailers and review platforms allow consumers to share their experiences with and opinions of a product in text reviews and to assign a rating (typically 1 to 5 stars) to the product. A higher rating indicates a more positive evaluation of the product. In order to help prospective consumers easily gauge product quality, retailers and review platforms commonly display the product's overall average rating along with other summarized rating statistics (e.g. the total number of ratings, the distribution of ratings) in more prominent places than individual reviews. Prior research found that aggregated rating statistics and individual reviews can each impact consumers' decision-making (e.g., Chevalier and Mayzlin 2006; Mudambi and Schuff 2010).

An implicit assumption among researchers and practitioners is that a product's overall average rating and other summary rating statistics matter more than individual reviews for predicting product sales. For example, large-scale meta-analyses of empirical evidence on determinants of product sales focused on the role of aggregated rating statistics rather than individual reviews (see Babić Rosario et al. 2016; Floyd et al. 2014). In practice, most online retailers and review sites display these aggregated rating statistics not only at the top of product review pages, but also at more prominent places (such as product listing pages) where no individual reviews are visible. In particular, the overall average rating is the most important cue for consumers to infer the quality of a product (De Langhe et al. 2015), because it is what consumers see first to form an initial impression of the product (Yin et al. 2016). Further, the overall average rating incorporates historical information from all the posted reviews and is therefore a more

comprehensive measure of product quality than individual reviews. This view is also in line with the anchoring effect, such that the average product rating can act as an anchor—a salient, initially presented value—before consumers make necessary adjustments based on individual reviews, and their final judgments would be close to the anchor (see Tversky and Kahneman 1974). Thus, this view predicts that the overall average rating of products is the primary driver of consumers' purchase decisions.

However, individual reviews can also affect consumers' purchase decisions. In addition to seeing aggregated rating statistics, consumers also consult individual reviews during the decision-making process. Prior literature studying individual reviews has examined the influence of rating and review text characteristics on the helpfulness of reviews because separating helpful from unhelpful reviews has both theoretical and practical implications (e.g., Mudambi and Schuff 2010; Yin et al. 2017; Yin et al. 2016). However, very few studies have connected individual reviews to consumers' intention to purchase the product or product sales (as an exception, see Shoham et al. 2017).

In addition, no research to our knowledge has directly compared the impact of aggregated rating statistics (derived from all historical ratings) with that of the most accessible individual reviews (those that appear on the first page of reviews) on product sales. In this paper, we revisit the commonly accepted, implicit assumption that a product's average rating matters more than individual reviews for product sales, and we explore whether and when consumers rely more on individual reviews in decision making. If individual reviews can easily sway product sales, then the current practice of review platforms, product manufacturers and retailers of highlighting the product's overall average rating and other summary rating statistics might not be impactful.

104

Drawing on the heuristics-and-biases literature (Tversky and Kahneman 1974), we propose two competing hypotheses on the relative impacts of a product's overall average rating and its most accessible reviews on consumers' purchase intention and product sales. On one hand, the *overall* average rating may matter more than individual reviews (namely the anchoring effect) because the product's average rating can serve as a salient anchor and bias subsequent consumer decision making toward the anchor. On the other hand, the average rating of the most accessible reviews may better predict product sales (namely the swaying effect) because prominently displayed individual reviews are more cognitively accessible (and thus more influential) for consumers. In addition, building on the metacognition literature (Smith and Swinyard 1988), we posit that a product's rating dispersion that shapes consumers' confidence in their initial beliefs about the product (Yin et al. 2016) can moderate the likelihood of anchoring and swaying effects. Utilizing two carefully designed experiments and an archival panel dataset collected daily from Apple's App Store over a two-month period, we find evidence supporting an overall swaying (versus anchoring) effect and the moderating role of products' rating dispersion.

Our study has a number of unique theoretical contributions. First, we challenge the widely accepted assumption that a product's overall average rating is the primary source of input in consumer decision making. Among the first attempts to compare the effects of overall average rating and easily accessible individual reviews, this paper provides both experimental and empirical evidence for a swaying effect—consumers' intention to purchase the product and product sales can be easily swayed by the individual reviews that are prominently displayed and most accessible for consumers.

Thus, the effect of supposedly all-encompassing cues such as the overall average rating on product sales may have been exaggerated in earlier research. Second, our study illustrates the potential and value of incorporating insights from the heuristics-and-biases literature into consumer behavior research. Our findings provide support for the availability heuristic but not the anchoring heuristic, suggesting that not all heuristics are equally applicable in the online reviews setting. Third, we extend the concept of belief confidence from the metacognition literature and identify rating distribution as a possible boundary condition for the anchoring and swaying effects. Together, these findings offer critical practical implications for product manufacturers, retailers, and review platforms that we explore later.

and empirical evidence for a swaying effect—consumers' intention to purchase the product and product sales can be easily swayed by the individual reviews that are prominently displayed and most accessible for consumers. Thus, the effect of supposedly all-encompassing cues such as the overall average rating on product sales may have been exaggerated in earlier research. Second, our study illustrates the potential and value of incorporating insights from the heuristics-and-biases literature into consumer behavior research. Our findings provide support for the availability heuristic but not the anchoring heuristic, suggesting that not all heuristics are equally applicable in the online reviews setting. Third, we extend the concept of belief confidence from the metacognition literature and identify rating distribution as a possible boundary condition for the anchoring and swaying effects. Together, these findings offer critical practical implications for product manufacturers, retailers, and review platforms that we explore later.

4.2 Hypotheses and Theory Development

With limited time and cognitive resources, consumers tend to utilize heuristics to navigate the vast amount of review information and to make decisions efficiently. People are generally "economy-minded" given their limited cognitive capacity, and they wish to achieve their goals in the most efficient ways possible with the least amount of effort (Chaiken 1980; Fiske and Taylor 1991). As a result, people generally prefer to use cognitive heuristics that enable them to make decisions efficiently. Heuristics are mental shortcuts or rules of thumb that most people rely on to reduce their effort and arrive at satisfactory decisions (Simon 1990). Although heuristics can be seen as adaptive, functional and even rational (Gigerenzer and Todd 1999), they can also lead to judgment biases and fallacies under certain situations. In particular, we focus on two of the primary heuristics from the heuristics-and-biases literature, the anchoring and availability heuristics (Tversky and Kahneman 1974). In addition, we propose a boundary condition for the effect of these two heuristics on consumers' decisions based on the notion of belief confidence (Smith and Swinyard 1988).

4.2.1 Anchoring Heuristic and Anchoring Effect

Anchoring heuristic is a robust and ubiquitous phenomenon in judgment and decision making (see Furnham and Boo 2011 for a comprehensive review). This heuristic suggests that people form judgments by first anchoring to a salient, initially presented value and then adjusting their evaluations from this starting point; as a result, the final judgments tend to be biased towards (or closer to) the anchor (Tversky and Kahneman 1974). This effect has been repeatedly and reliably demonstrated in various contexts,

including valuation and purchasing decisions (Ariely et al. 2003; Mussweiler et al. 2000; Wansink et al. 1998).

The dominant explanation for the anchoring effect is confirmatory hypothesis testing. This explanation argues that decision makers consider the initially presented anchor as a plausible value or hypothesis to be tested out (Strack and Mussweiler 1997). To increase the credibility of a hypothesis and defend their initial beliefs, people have a common tendency to search for information that confirms their first impression or hypothesis (Nickerson 1998). Their preference for such anchor-consistent information will increase the validity and dependence of the anchor value, which in turn leads the final judgment to be closer to the anchor (Mussweiler and Strack 1999). This view has been empirically supported by numerous studies that vary external anchors given by the researchers (e.g., Chapman and Johnson 1999; Mussweiler and Strack 2001; Wegener et al. 2010). For instance, the anchoring effect was observed when participants in an experiment were asked to decide a minimum selling price for a lottery; the selling price was influenced by a random anchor derived from their social security number and provided to the decision makers (Chapman and Johnson 1999). In addition, the anchoring effect has also been demonstrated when the anchors are integral to the decision-making context and socially derived, such as the consensus or average value of others' decision making that is publicly available and relevant (Meub and Proeger 2015; Phillips and Menkhaus 2010). For example, in an estimation task, participants were asked to predict future values of a simple formula; they were provided with the estimations from all other participants, and their predications were biased toward this socially derived anchor (Meub and Proeger 2015).

In our context, a product's overall average rating consolidated from all the historical opinions from its prior customers serves as a salient anchor for prospective consumers (Sun 2012). This aggregated, product-level information cue is often among the first that consumers see on review websites about a product. In Apple's App Store, for example, when consumers browse the top-ranking charts or search for apps, the title of each app in the list is displayed along with its overall average rating. Within a particular app, the overall average rating of the app is presented twice—at the top and bottom of the description section—before consumers can scroll to read individual reviews. The overall average rating of a product is the primary source of input for consumers to form their initial beliefs about the product's quality (De Langhe et al. 2015). Extending the reasoning of anchoring heuristic to our context, after consumers form initial beliefs based on the anchor—overall average rating, they should favor anchor-consistent information to support and defend their initial evaluations of the product. Supporting this confirmatory hypothesis testing explanation of the anchoring effect, recent investigations into the impact of a product's overall average rating (as a salient anchor) also provided empirical evidence for a confirmation bias—consumers evaluate confirmatory reviews in line with the anchor more favorably (Baek et al. 2015; Yin et al. 2016). The greater perceived value of reviews confirming the anchor increases the likelihood that consumers' intention to purchase the product will be determined more by the anchor than by individual reviews. Taken together, we propose the first hypothesis below.

Hypothesis 1a. The overall average rating of a product has an anchoring effect; that is, consumers' intention to purchase the product is influenced more by the product's

overall average rating that incorporates information from all historical reviews than by

the average rating of the few reviews that consumers observe.

4.2.2   *Availability Heuristic and Swaying Effect*

In contrast to the anchoring heuristic, the availability heuristic offers a different

viewpoint and predicts the opposite to Hypothesis 1a. As another prominent heuristic

from the heuristics-and-biases literature, availability heuristic states that decision makers

rely on accessible instances or immediate examples of a target object that easily come to

mind as the basis for their decision making (Tversky and Kahneman 1973). Decision

makers estimate the probability of a target object by assessing its accessibility because of

their natural co-occurrences (e.g., more frequent examples are recalled better and faster

than less frequent examples, and more likely instances are easier to imagine than unlikely

instances) (Tversky and Kahneman 1973). Based on the "ease of retrieval" explanation,

people often use the ease with which examples can be brought to mind to infer the

importance, frequency, and typicality of such examples (Schwarz et al. 1991). As a result,

decision makers' judgments are biased towards easy-to-retrieve and accessible instances.

While a variety of factors can contribute to the availability or accessibility of

information, vividness is considered the most important factor (Nisbett and Ross 1980).

Vividness of a piece of information is defined as the extent to which that information is

concrete and image provoking (Nisbett and Ross 1980). Experimental studies examining

vividness has manipulated this factor through narrative information (such as verbal

descriptions) versus statistical information (such as summarized statistics) (e.g., Keller

and Block 1997). For instance, in a study manipulating the vividness of the product

110

depiction through narrative versus statistical information, participants were presented with easy-to-imagine verbal product descriptions or summarized numerical ratings (Petrova and Cialdini 2005). Because narrative information is easier to imagine, visualize and recall than statistical, numerical information, the former comes to mind more easily and affects decisions to a greater extent than the latter (Petrova and Cialdini 2005).

Because of this availability heuristic, consumers in our context may rely more on individual reviews they observe that are more concrete and vivid than a product's overall average rating. An individual review contains detailed, concrete experiences and opinions from a previous consumer, while the overall average rating is a consensus evaluation abstracted from all prior reviews of the product. Because individual reviews are more concrete and accessible than the overall average rating, the former should be easier to recall than the latter when consumers make purchase decisions. Based on the availability heuristic and "ease of retrieval" explanation, individual reviews should be perceived as more important and more typical of a likely consumption experience with the product. As a result, individual reviews should play a greater role in consumers' purchase decision. Thus, we propose the following competing hypothesis.

Hypothesis 1b. Individual reviews of a product have a swaying effect; that is, consumers' intention to purchase the product is influenced more by the average rating of the few reviews that consumers observe than by the product's overall average rating that incorporates information from all historical reviews.

4.2.3   Boundary Condition: Confidence in Initial Beliefs

The likelihood and extent of anchoring versus swaying effects proposed in H1a and H1b may depend on consumers' confidence in their initial beliefs about the product. A product's overall average rating incorporates the evaluations from all previous customers, and consumers rely on this signal to form their initial beliefs about the quality of the product (De Langhe et al. 2015). In addition, consumers may have different levels of confidence in their initial beliefs (Petty et al. 2007), and the distribution of a product's ratings can shape consumers' confidence in their beliefs (Yin et al. 2016). Confidence in initial beliefs (or belief confidence for short) refers to consumers' perceived confidence and certainty with which they consider their initial beliefs to be correct (Smith and Swinyard 1988). Confidence is a fundamental dimension of metacognition (or secondary cognition) because it involves second-order thoughts (e.g., is my evaluation accurate or valid?) about first-order thoughts (e.g., how positive or negative is my evaluation toward the product?) (Petty et al. 2007). Metacognition can strengthen or weaken first-order thoughts, but the online word-of-mouth literature has just started investigating its role in consumer judgment and decision making (Yin et al. 2016).

We first propose that consumers' confidence in initial beliefs could influence the likelihood and extent of the anchoring effect. As we argued earlier, the anchoring effect arises because the anchor (i.e., a product's overall average rating) shapes consumers' initial beliefs about the product (Sun 2012) and that consumers prefer individual reviews confirming these initial beliefs. When consumers are more confident and certain about their initial beliefs, they should perceive the anchor to be more accurate and dependable. As suggested by the anchoring heuristic, after consumers form their initial beliefs based

112

on the anchor, they may adjust their evaluations based on new information (Tversky and Kahneman 1974). However, this adjustment is less likely and tends to be smaller when consumers have higher confidence in the anchor and deem the anchor to be more dependable than new information (Mussweiler and Strack 1999; Mussweiler and Strack 2000). As a result, consumers' final decisions should be closer to the anchor (i.e., stronger anchoring effect) when they are more confident about their initial beliefs.

On review sites, the distribution of ratings is a prominent signal that can inform consumers about the accuracy and reliability of the overall average rating, and thus can influence the anchoring effect. After forming their initial beliefs about a product based on its overall average rating, consumers can easily validate the accuracy of their initial beliefs by consulting the dispersion of rating distributions. Specifically, low dispersion indicates a high level of consensus that the opinions from prior customers are highly consistent with each other, whereas high dispersion indicates low consensus (Moe and Trusov 2011). In addition, lower dispersion can increase consumers' confidence and certainty in their initial beliefs because it indicates that the initial belief based on the anchor is more accurate and reflective of prior customers' opinions (Petrocelli et al. 2007a; Yin et al. 2016). Following earlier reasoning, consumers in this situation are less likely to make adjustments based on new information provided in individual reviews. Therefore, their final decisions are less likely to depart from the anchor, and we propose the second hypothesis below.

Hypothesis 2: The anchoring effect (i.e., influence of a product's overall average rating) is stronger for products that have a lower rating dispersion.

In addition, the likelihood and extent of the swaying effect may also depend on consumers' confidence in their initial beliefs. As we discussed earlier, the swaying effect could be activated because vivid and concrete individual reviews are easier to retrieve and that consumers consider these easy-to-imagine reviews to be more important when making decisions (Schwarz et al. 1991). When consumers are less confident and certain about their initial beliefs, they should look for and rely more on the new information to reduce uncertainty (e.g., Bauer 1967; Elliot 1999). As they pay more attention to the new information provided in individual reviews, the swaying effect should be stronger because consumers are more likely to consider vivid and concrete verbal descriptions from individual reviews (compared with a product's overall average rating) and deem such information as more important in this situation.

On review sites, the distribution of ratings can change the swaying effect by changing consumers' confidence in their initial beliefs. Because high dispersion indicates low consensus among prior customers (Moe and Trusov 2011), consumers are less confident in the accuracy of a product's overall average rating, and they pay more attention to individual reviews instead. Following earlier arguments, consumers in this situation are more likely to recall vivid and concrete information from individual reviews when they make purchase decisions, resulting in a stronger swaying effect. Taken together, we propose the third hypothesis below.

Hypothesis 3: The swaying effect (i.e., influence of the average rating of the few reviews that consumers observe) is stronger for products that have a higher rating dispersion.

To test these hypotheses, we conducted three studies using experimental and econometrics methods. In Study 1, we conducted an experiment to explore the existence of anchoring versus swaying effects, and directly manipulated the overall average rating of two products and their most recent reviews. Study 2 extended our manipulation to the dispersion of rating distributions and explored the boundary condition of the main finding in Study 1. In Study 3, we utilized a unique panel dataset collected daily from Apples' App Store over a two-month period and replicated the findings of the two earlier studies.

4.3 Study 1

In the first study, we designed an experiment to test the competing hypotheses. Specifically, in a hypothetical online decision-making task, participants read the rating profiles (i.e., summary statistics of user ratings) and the 3 most recent individual reviews of two products, and then make purchase decisions between the two products. We varied the overall average rating (4 stars vs. 4.5 stars) and the ratings of the individual reviews (slightly positive vs. slightly negative) in such a way that one product would be superior based on overall average rating while the other product would be superior based on individual reviews. Consumers' purchase intentions and choice between the two products would indicate the likelihood and extent of anchoring vs. swaying effects.

4.3.1   Stimulus Materials

We used the digital camera as our context because it is a familiar product for most people. A digital camera is a camera that captures and stores photographic images in digital memory. To vary the valence of 3 individual reviews of each product, we developed 3 sets of treatment reviews (with a positive version and a negative version in each set) for each of the two products. To remove possible confounds, we also conducted a pretest to ensure that two versions within each review set are equivalent in extremity, and that different review sets are equivalent with regard to information quantity, concreteness, extremity, helpfulness, emotional intensity, realism, and reading difficulty.

We started with three common camera features—ease of use, LCD, and image stabilization—and then wrote 4 sets of individual reviews for each feature by consulting actual camera reviews from Amazon. Within each review set, we first prepared a positive version, and then constructed a corresponding negative version by adding negations and using antonyms while holding the substantial content identical. Because we also kept the number of words in each review at around 25, the only difference between the positive and negative versions within each set is the valence.

In a pretest, we recruited 55 participants from Amazon Mechanical Turk (MTurk), and asked each participant to read and evaluate 12 reviews, one review at a time. We randomly assigned them to read one version (either positive or negative) of the reviews in each review set. After reading each review, we asked each participant to report their evaluations of the review's 1) extremity (e.g., "not at all positive / very positive"), 2) information quantity (e.g., "contains very little information / contains a great deal of

information"), 3) concreteness (e.g., "not at all concrete / very concrete"), 4) helpfulness (e.g., "not at all helpful / very helpful"), 5) emotional intensity (e.g., "contains little emotion / contains a great deal of emotion"), 6) realism (e.g., "not at all realistic / very realistic"), and 7) reading difficulty (e.g., "very hard to read / very easy to read"). Each variable was measured by two items adapted from prior literature and along a 9-point scale (see Appendix H for all the measures and their sources). We next conducted independent-samples t-tests of extremity between the positive and negative versions in each review set and paired-samples t-tests of all other variables (e.g., information quantity, concreteness, extremity, etc.) across different review sets. Based on the pretest results, we selected 3 sets of reviews for each of the two treatment products that satisfy our criteria (see Table 4-1): each review set describes a different feature; comparisons in extremity of the positive and negative versions within each set yielded a t-value of at most 1.589 with a p-value of at least 0.118; comparisons in all other relevant variables of all the positive (negative) versions across different review sets yielded a t-value of at most 1.344 (1.705) with a p-value of at least 0.191 (0.100). Therefore, the 3 sets of treatment reviews we chose for each product are not significantly different in extremity between two versions within each review set, or in other relevant variables (e.g., information quantity, concreteness, extremity, etc.) across different review sets.

Table 4-1 3 Sets of Reviews for Each Product.

Set #	Product 1	
	Positive Version	**Negative Version**
1	This camera is user-friendly compared to other entry level cameras. After just a few days of use, I found it really intuitive.	This camera is not user-friendly compared to other entry level cameras. After just a few days of use, I found it really complicated.
2	The camera has an excellent LCD screen, which is large enough for most people. It works well in both high and low light conditions.	The camera has a poor LCD screen, which is not large enough for most people. It doesn't work well in either high or low light condition.
3	The image stabilization works as expected. It can correct the impact of minor accidental hand motion. The pictures taken in unsteady situations are clear.	The image stabilization doesn't work as expected. It fails to correct the impact of accidental hands motion. The pictures taken in unsteady situations are blurry.

Set #	Product 2	
	Positive Version	**Negative Version**
1	Everything on this camera is easy to use. It's a good choice for people who don't have much experience with digital cameras.	Everything on this camera is hard to use. It's a bad choice for people who don't have much experience with digital cameras.
2	The LCD screen has quick feedback. After I take shots, I can see the pictures flashed onto the LCD screen instantly.	The LCD screen has slow feedback. After I take shots, I can see the pictures flashed onto the LCD screen after a while.
3	The image stabilization is effective. This feature ensures the clarity of photo if the camera is slightly moved when I take the photo.	The image stabilization is not effective. This feature cannot ensure the clarity of photo if the camera is slightly moved when I take the photo.

4.3.2   Procedure

53 undergraduate students from a U.S. university took part in the experiment in exchange for extra credit. In the cover story, participants were asked to imagine that they were planning to purchase a digital camera from Amazon.com, and their search returned two different digital cameras with the same price of $549.99. Then they were asked to read

the rating profile (including the average rating and the number of ratings from prior users) and the 3 most recent reviews displayed on the first review page for both cameras, one product at a time, with the order of the two products and the order of each product's 3 reviews counterbalanced.

Both cameras had accumulated hundreds of reviews, but had different average rating scores of 4 and 4.5 (out of 5) stars. In addition, we varied the ratings of the three most recent reviews in such a way that the product with a lower overall average rating would be superior based on individual reviews, and vice versa. Each product is better based on either the overall average rating or the most recent reviews, but not both; this design allowed us to test the likelihood and extent of anchoring versus swaying effects: if participants have higher intention to purchase the 4.5-star product and choose it over the 4-star product, then their purchase decisions are influenced more by the overall average rating, indicating an anchoring effect; if they have higher intention to purchase the 4-star product (whose individual reviews are more positive) and choose it over the 4.5-star product, then they are swayed by individual reviews when making purchase decisions. Specifically, participants were presented with 2 positive and 1 negative reviews for the 4-star product, while they were presented with 1 positive and 2 negative reviews for the 4.5-star product. To strengthen the manipulation of the review valence, we also displayed the rating score (5 stars for the positive review and 1 star for the negative review) along with the review content.

After observing the rating profile and reading the 3 most recent individual reviews of each product, participants were asked to report their intention to purchase the product using a 9-point scale adapted from Dodds et al. (1991) and Goldberg and Gorn (1987) (e.g.,

119

"If you were thinking of buying a digital camera, how likely is it that you would buy Camera Model A?"). Then participants were presented with the rating profiles and individual reviews of two products side by side (see Figure 4-1 for a screenshot in one condition), with the first product they evaluated earlier appearing on the left side of the screen. The product on the left could have an overall average rating of either 4 or 4.5 stars (due to the counterbalancing of the two products' order), thus mitigating a potential confound of the location effect (e.g., 4.5-star product being displayed on the left or right). After observing the rating profiles and individual reviews of two products, participants were asked to choose one camera between two options for purchase, using an 8-point scale (1 = "definitely choose Camera A", 8 = "definitely choose Camera B"). We used an 8-point scale here so that participants cannot keep a neutral stance between the two options. As a manipulation check, participants were also asked to recall the average rating of each product, and to evaluate the valence of the 3 individual reviews as a whole for each product (see Appendix I for all measures used in this study).
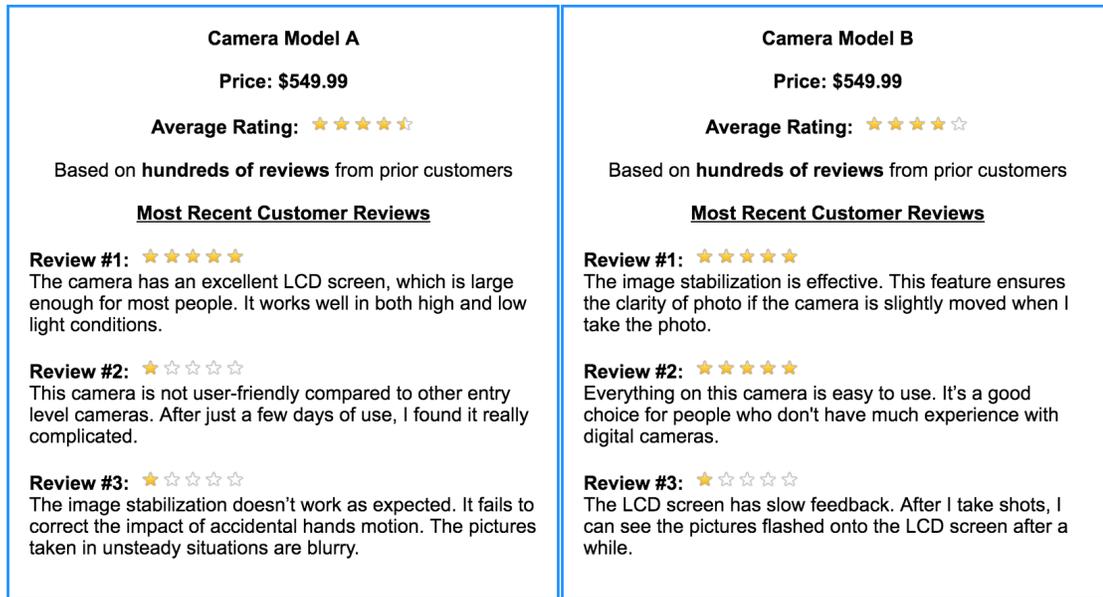
Figure 4-1 Rating Profiles and Individual Reviews of Two Camera Options.

4.3.3 Results

We first conducted manipulation checks for the two variables we manipulated in the study. Although not all participants were able to correctly recall the overall average rating of the two products, the mean of their recalled overall average rating of the 4-star product was significantly lower than that of the 4.5-star product ($M = 3.84$ vs. 4.05, $F(1, 49) = 4.831$, $p = 0.033$), indicating that our manipulation of the products' overall average rating was successful and in expected direction. In addition, the perceived valence of the most recent reviews of the 4-star product (2 positive and 1 negative reviews) was significantly more positive than that of the 4.5-star product (1 positive and 2 negative reviews) ($M = 6.63$ vs. 3.49, $F(1, 49) = 112.520$, $p < 0.001$), indicating that our manipulation of the valence of the most recent reviews was also successful.

121

To explore the relative impacts of the products' overall average rating and most recent reviews on consumers' purchase intention, we conducted a repeated-measure ANCOVA with the two products entered as a within-subject factor and treatment order entered as a covariate. Consistent with a swaying effect, results revealed that participants' intention to purchase the 4-star product was significantly higher than their intention to purchase the 4.5-star product ($M = 6.15$ vs. 3.59, $F(1, 49) = 72.272$, $p < 0.001$). Thus, among the two competing hypotheses (H1a vs. H1b), H1b is supported.

In addition, we investigated whether the swaying effect also shaped participant choice. Participants provided their choice between the two product options along an 8-point scale (1 = "definitely choose Camera Model A," 8 = "definitely choose Camera Model B"). We re-coded the choice values so that a lower value indicates participants' preference for the 4.5-star product and that a higher value indicates their preference for the 4-star product. Then we conducted a one-sample t-test to compare the mean of participants' choices with the midpoint (4.5) of the scale. Results revealed that the mean value of recoded responses was 6.32, which was significantly above the midpoint ($t(52) = 8.970$, $p < 0.001$). Thus, participants preferred to choose the 4-star product compared with the 4.5-star product, providing additional evidence for H1b and a swaying effect.

4.3.4 Discussion

In Study 1, we examined the likelihood of anchoring versus swaying effects by manipulating two products' overall average rating and their most recent reviews. The results provided evidence for the swaying effect of individual reviews proposed in H1b:

consumers' intention to purchase a product is influenced more by the most accessible reviews than by the products' overall average rating.

Our design in Study 1 had two notable limitations. First, we focused on testing H1a versus H1b, but were not able to test the moderating role of rating dispersion in Study 1. Second, the overall average ratings of two products utilized in this study were very positive (i.e., 4 and 4.5 stars), with a tiny difference of 0.5 star. Participants might be swayed by individual reviews because the overall average ratings we manipulated were too positive and differing too little (although 0.5-star difference based on hundreds of individual reviews is still a big difference at the product level). Therefore, we designed another experiment to address these limitations.

4.4   Study 2

In Study 2, we utilized the similar design as Study 1 with a few exceptions. First, we manipulated the overall average ratings of the two products to be less positive with a greater difference (2.5 vs. 3.5 stars). Second, we explored the boundary condition of the anchoring versus swaying effects by manipulating the rating dispersion between-subjects at two levels.

4.4.1   Procedure and Measures

161 undergraduate students from a U.S. university took part in this experiment in exchange for extra credit. This study followed the similar procedure as in Study 1, except that each subject was randomly assigned to the low or high rating dispersion condition. Those in the low rating dispersion condition saw "90% of the reviewers rated the camera

at 3 and 2 (4 and 3) stars" at the bottom of each product's rating profile, while those in the high rating dispersion condition saw "10% of the reviewers rated the camera at 3 and 2 (4 and 3) stars". An example of all the information for both products in the high dispersion condition is presented in Figure 4-2. As a manipulation check, we also asked participants to report the rating dispersion of each product with two items adapted from He and Bond (2015). See Appendix I for all the measures used in this study.

Camera Model A	Camera Model B
Price: $549.99	**Price: $549.99**
Average Rating: ★★⯪☆☆	**Average Rating:** ★★★⯪☆
Based on **hundreds of reviews** from prior customers	Based on **hundreds of reviews** from prior customers
10% of the reviewers rated the camera at **3 and 2 stars**	**10%** of the reviewers rated the camera at **4 and 3 stars**
Most Recent Customer Reviews	**Most Recent Customer Reviews**
Review #1: ★★★★★ The camera has an excellent LCD screen, which is large enough for most people. It works well in both high and low light conditions.	**Review #1:** ★☆☆☆☆ Everything on this camera is hard to use. It's a bad choice for people who don't have much experience with digital cameras.
Review #2: ★☆☆☆☆ The image stabilization doesn't work as expected. It fails to correct the impact of accidental hands motion. The pictures taken in unsteady situations are blurry.	**Review #2:** ★☆☆☆☆ The LCD screen has slow feedback. After I take shots, I can see the pictures flashed onto the LCD screen after a while.
Review #3: ★★★★★ This camera is user-friendly compared to other entry level cameras. After just a few days of use, I found it really intuitive.	**Review #3:** ★★★★★ The image stabilization is effective. This feature ensures the clarity of photo if the camera is slightly moved when I take the photo.

Figure 4-2 Rating Profiles and Individual Reviews of Two Camera Options (High Dispersion Condition).

First, we conducted manipulation checks of the overall average rating, the perceived valence of the most recent reviews, and the rating dispersion. Results revealed that the recalled overall average rating of the 2.5-star product was significantly lower than that of the 3.5-star product ($M = 2.86$ vs. 3.22, $F(1, 157) = 26.142$, $p < 0.001$), and that the perceived valance of 3 reviews in the "2 positive and 1 negative reviews" condition was significantly higher than that in the "1 positive and 2 negative reviews" condition ($M = 6.36$ vs. 3.54, $F(1, 157) = 415.875$, $p < 0.001$). In addition, the perceived rating dispersion in the low dispersion condition was significantly lower than that in the high dispersion condition for the 2.5-star product ($M = 4.14$ vs. 5.66, $F(1, 156) = 21.656$, $p < 0.001$), and the same pattern applied for the 3.5-star product ($M = 4.05$ vs. 6.18, $F(1, 156) = 42.867$, $p < 0.001$). Therefore, the manipulations of all variables were deemed successful.

Next, to explore the likelihood of anchoring versus swaying effects as well as the moderating role of rating dispersion, we conducted ANCOVA with participants' intention to purchase the product entered as the dependent variable, the two products entered as a within-subject factor, rating dispersion entered as a between-subjects factor, and treatment order entered as a covariate. Results revealed that participants' intention to purchase the 2.5-star product was significantly higher than their intention to purchase the 3.5-star product ($M = 4.60$ vs. 3.91, $F(1, 156) = 9.707$, $p = 0.002$), providing additional evidence for the swaying effect. Therefore, H1b is supported. In addition, the interaction between the swaying effect of individual reviews and rating dispersion was marginally significant ($F(1, 156) = 3.331$, $p = 0.070$). Pairwise comparisons revealed that when the rating dispersion was high, participants' intention to purchase the 2.5-star product was

significantly higher than their intention to purchase the 3.5-star product ($M = 4.89$ vs. 3.81, $F(1, 156) = 12.280$, $p = 0.001$). However, when the rating dispersion was low, the difference in purchase intention was not significant ($F(1, 156) = 0.828$, $p = 0.364$). Therefore, these results provided evidence for the moderating impact of rating dispersion on the swaying effect as we hypothesized in H3.

In addition, we investigated the moderating role of rating dispersion when participants made a choice between the two product options along an 8-point scale (1 = "definitely choose Camera Model A," 8 = "definitely choose Camera Model B"). Following the similar analysis in Study 1, we recoded participants' choice, with a value above the midpoint (4.5) indicating a preference for 2.5-star product and a value below the midpoint indicating a preference for 3.5-star product. We conducted a one-sample t-test for high and low dispersion conditions respectively. When the rating dispersion was high, the mean value of consumers' choice ($M = 4.94$) was marginally significantly above the midpoint ($t(80) = 1.762$, $p = 0.082$), indicating that participants preferred the 2.5-star product to the 3.5-star product in this condition. However, when the rating dispersion was low, the mean value of consumers' choice ($M = 4.21$) was not significantly below the midpoint ($t(79) = 1.090$, $p = 0.279$). Hence, the results of one-sample t-tests provided additional evidence for H3. In addition, we conducted ANCOVA with consumers' choice versus 4.5 (the midpoint of the scale) entered as a within-subject factor, rating dispersion entered as a between-subjects factor, and treatment order entered as a covariate. Results revealed that the interaction between the within-subject and between-subjects factors was significant ($F(1, 156) = 3.331$, $p = 0.048$). Pairwise comparisons of the mean value of

consumers' choice versus the midpoint under low and high rating dispersion conditions were consistent with the results of one-sample t-tests reported earlier.

4.4.3 Discussion

In Study 2, we examined the likelihood of anchoring versus swaying effects as well as the moderating role of the rating dispersion for products whose overall average ratings were closer to neutral and more different from each other. The study not only replicated Study 1's findings of the swaying effect (H1b), but also provided evidence for a boundary condition of the swaying effect as hypothesized in H3.

The design of the first two experiments allowed us to provide causal evidence for the swaying effect and its boundary condition, as we can effectively rule out potential confounds by directly manipulating the overall average rating and the most recent reviews while keeping all other variables identical. However, the experimental manipulation has artificial elements. For example, we manipulated rating dispersion using words (10% vs. 90%) rather than a distribution graph commonly used in real review sites because it is nearly impossible to vary both the average and distribution of two products' ratings while holding all other aspects of the rating graphs identical. In addition, our experiments cannot test the boundary condition of the anchoring effect proposed in H2, as both studies found an overall swaying effect. In order to address these concerns, we conducted a final study using actual rating and reviews collected from Apple's App Store.

4.5   Study 3

The primary goal of this study was to test our proposed hypotheses in a real-word setting with actual ratings and reviews of apps from Apple's App Store. Existing users of an app can evaluate the app by assigning a rating on a scale of 1 to 5 stars. In addition, users can provide a detailed description of their experiences with the app in a text review. Users can also submit a rating without writing a text review. When prospective consumers read the text reviews of an app, they can indicate whether they find a review "helpful" or "not helpful" by clicking on "Yes" or "No" buttons next to the review. By default, the review section of an app displays 10 reviews per page, sorted by review helpfulness.

4.5.1   Data

We collected daily reviews of apps from Apple's App Store between July 1st and August 31st, 2013. We targeted 538 apps that ranked in the top 100 in Apple's App Store (based on downloads) at least once during June 2013. Apple classifies these apps into 21 categories (such as games, business, finance, and news). Because the 10 reviews on the first review page of an app are the most accessible for prospective consumers, we extracted the rating and content of the first 10 reviews of each app on each day. Since we did not change the default sort order of reviews in our data collection, it is highly likely that most consumers saw the same 10 reviews of an app on the first page. For each app, we also tracked the following app-level data that changed over time: the overall ranking of the app, the average rating of the app, the total number of ratings of the app, the distribution of the ratings for the app (e.g., number of one-star ratings, number of two-star

ratings, etc.), the price of the app, whether the app released an updated version on a specific date, and the date when the app was first launched. Our final sample contained data of 482 (out of 538) apps during the study period of 62 days that had sufficient information to calculate the variables we describe next.

4.5.2   Variable Definitions

Table 4-2 shows the definitions of the variables used in the empirical analysis, while Tables 4-3 and 4-4 show statistics and correlations for these variables. For the variable definitions in Table 4-2, i indexes an app and t indexes the event time (day) during our study period. Our dependent variable is $ARank_{it}$, the overall rank of app i at time t. It was calculated based on the number of downloads of the app and is displayed in the "Top Charts" list made available by Apple. We collected the overall rank of each app for each day in the study period from Apple's App Store. The overall rank ($ARank_{it}$) is a proxy for product sales (and an extension of consumers' purchase intention) in the online app context. A smaller numeric rank indicates a greater number of downloads (product sales).

Table 4-2 Variable Definitions.

Variable Name	Operationalization
$ARank_{it}$	Rank of app i at time t based on the number of downloads
$ARating_{it}$	Average rating of app i at time t
$RARating1st10_{it}$	Average rating of the first 10 reviews of app i at time t
$APrice_{it}$	Price per download of app i at time t
$AUpd_{it}$	= 1 if app i released an update (new version) at time t, 0 otherwise
$ARCount_{it}$	Cumulative number of ratings for app i at time t
$ADispersion_{it}$	Standard deviation of the ratings for app i at time t
$ADays_{it}$	Age of app i (in days) at time t
$RALength1st10_{it}$	Average number of words in the first 10 reviews of app i at time t

Table 4-3 Descriptive Statistics.

Variable	N	Mean	Std Dev	Min	Max
$ARank_{it}$	25358	354.84	401.67	1	1500
$ARating_{it}$	32065	4.18	0.61	1	5
$RARating1st10_{it}$	30094	3.83	1.00	1	5
$APrice_{it}$	33356	1.23	4.36	0	69.99
$AUpd_{it}$	33356	0.02	0.15	0	1
$ARCount_{it}$	33356	2700	7653.04	0	104407
$ADispersion_{it}$	32065	1.15	0.35	0	2
$ADays_{it}$	33356	499.82	501.20	3	1877
$RALength1st10_{it}$	30094	16.25	4.43	0.1	21.5

Table 4-4 Correlations.

Variable	1	2	3	4	5	6	7	8	9
1 $ARank_{it}$	1								
2 $ARating_{it}$	-0.02	1							
3 $RARating1st10_{it}$	0.03	0.58	1						
4 $APrice_{it}$	0.03	0.13	0.10	1					
5 $AUpd_{it}$	-0.01	-0.03	0.01	-0.02	1				
6 $ARCount_{it}$	-0.15	0.17	-0.12	-0.05	-0.04	1			
7 $ADispersion_{it}$	0.00	-0.79	-0.54	-0.10	0.02	-0.18	1		
8 $ADays_{it}$	-0.25	-0.07	-0.07	0.02	-0.01	0.04	0.04	1	
9 $RALength1st10_{it}$	-0.08	-0.17	-0.31	0.11	-0.09	-0.04	0.22	0.12	1

Our primary independent variables are the overall average rating of an app and the average rating of the most accessible reviews (that appear on the first page of reviews). $ARating_{it}$ is the average rating of app i at time t based on all the ratings provided by consumers for the latest version of the app. The App Store site displays the average rating of each app prominently along with other summarized statistics (such as the number of ratings), so we could obtain $ARating_{it}$ directly from the App Store. To represent the average rating of the most accessible reviews, we created one variable $RARating1st10_{it}$ that is the average rating of the 10 reviews displayed on the first review page of app i at time t.

Our moderator is the dispersion of the ratings, $ADispersion_{it}$, measured by the population standard deviation of all the review ratings from previous users for app i at

time t.[6] We also controlled for several app-date level variables in our analysis. $APrice_{it}$

is the price of downloading app i at time t. The price of an app may change over time due

to a promotion or other reasons, and the price could affect consumers' download

decisions. $AUpd_{it}$ is an indicator which equals 1 if app i released an update (new version)

at time t, and 0 otherwise. Consumers' intention to purchase or download an app can be

influenced by whether the app has an updated version, because a new version often

contains improvements and there may be promotions associated with a new version.

$ARCount_{it}$ is the cumulative number of ratings for app i at time t. The number of ratings

of an app can affect consumers' purchase decisions as well since it indicates the

prominence of the app. $ADays_{it}$ indicates the number of days at time t since the app was

first launched. Finally, we controlled for the average number of words in the 10 reviews

on the first review page of app i at time t through the variable $RALength1st10_{it}$.

4.5.3 Methods and Empirical Analysis

We evaluated the following two fixed effects panel data models in Stata.

Following prior literature (Chevalier and Mayzlin 2006; Forman et al. 2008), we log

transformed the dependent variable to account for scale effects (ranks of popular apps

may change more) and ease interpretation (the coefficients approximately indicate a

percentage change in rank). In the models, U_i is the fixed effect intercept for app i and $\in_{it}$

is the error term.

---

[6] All of the following results were consistent when we used an alternative measure of the rating dispersion, measured by the sample standard deviation.

There are two main sources of endogeneity possible in our analysis. First, unobserved characteristics of an app (such as its quality) can affect both its sales rank (dependent variable) and its ratings (independent variables). Second, an app's sales rank can affect how consumers rate the app (reverse causality) because consumers may rate popular apps more highly. Consequently, we identify the effects of $ARating_{it}$ and $RARating1st10_{it}$ on our dependent variable ($ARank_{it}$) through the following characteristics of our model. First, time invariant and unobserved characteristics of an app are captured through the fixed effects intercept term in the model. Second, note that the forward period dependent variable ($ARank_{it+1}$) in the two models below reduces identification problems associated with reverse causality between the dependent and independent variables. Third, we focused on apps that had a minimum number of ratings (e.g., 100, 150, and 200) on each day during the study period because this lower bound reduced multi-collinearity between the overall average rating and the average rating of the 10 reviews displayed in the first review page when the app has only a few reviews.[7] This ensured that the overall average rating of an app differed from the average rating of the 10 reviews on the first review page, making it possible to evaluate accurately the impact of each. Further, as a robustness check, we also restrict the analysis to those apps that had a variation in the average rating over time, and we explain this later in the paper.

---

[7] Results presented in the left panel of Table 4-5 are based on apps that received at least 150 ratings on any day. Results were consistent when the lower bound of the number of ratings was set to 100 and 200.

$$Ln(ARank_{it+1}) = \beta_0 + \beta_1 APrice_{it} + \beta_2 AUpd_{it} + \beta_3 ARCount_{it} + \beta_4 ADays_{it} + \beta_5 ARating_{it} \quad (1)$$

$$+ \beta_6 RALength1st10_{it} + \beta_7 RARating1st10_{it} + U_i + \epsilon_{it}$$

$$Ln(ARank_{it+1}) = \beta_0 + \beta_1 APrice_{it} + \beta_2 AUpd_{it} + \beta_3 ARCount_{it} + \beta_4 ADays_{it} + \beta_5 ARating_{it} \quad (2)$$

$$+ \beta_6 RALength1st10_{it} + \beta_7 RARating1st10_{it} + \beta_8 ADispersion_{it} + \beta_9 ARating_{it} * ADispersion_{it}$$

$$+ \beta_{10} RARating1st10_{it} * ADispersion_{it} + U_i + \epsilon_{it}$$

The results shown in Table 4-5 are based on 482 apps in our sample. The left panel shows the coefficient estimates from the main models. Model (1) includes the control variables, the overall average rating ($ARating_{it}$), and the average rating of the ten reviews that appear on the first review page ($RARating1st10_{it}$). The coefficient of the $ARating_{it}$ variable is not significant, while the coefficient for the $RARating1st10_{it}$ variable is negative and significant ($\beta = -0.042$, $p < 0.01$). The latter indicates that each unit increase in the average rating of the ten reviews on an app's first review page improves the app's ranking by 4.2%. The estimated coefficient of $RARating1st10_{it}$ is almost three times that of $ARating_{it}$. Thus, the results from Model (1) support the swaying effect proposed in Hypothesis H1b and not Hypothesis H1a.

Table 4-5 Fixed Effect Models.

DV: Ln($ARank_{it+1}$)	Main Models		Robustness Check	
	Model 1	**Model 2**	**Model 3**	**Model 4**
$APrice_{it}$	0.519***	0.515***	0.205***	0.198***
	(0.014)	(0.014)	(0.015)	(0.015)
$AUpd_{it}$	-0.180***	-0.175***	-0.062*	-0.070*
	(0.044)	(0.043)	(0.036)	(0.036)
$ARCount_{it}$	0.000***	0.000***	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
$ADays_{it}$	0.004***	0.004***	0.007***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)
$ARating_{it}$	0.014	-0.543***	0.047**	-0.279***
	(0.021)	(0.088)	(0.019)	(0.070)
$RALength1st10_{it}$	0.005***	0.006***	0.001	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
$RARating1st10_{it}$	-0.042***	0.021	-0.061***	0.035
	(0.009)	(0.034)	(0.011)	(0.050)
$ADispersion_{it}$		-1.545***		-0.719***
		(0.244)		(0.162)
$ARating_{it}$ * $ADispersion_{it}$		0.307***		0.169***
		(0.060)		(0.047)
$RARating1st10_{it}$ * $ADispersion_{it}$		-0.063**		-0.074**
		(0.027)		(0.034)
$Constant$	1.665***	4.302***	0.653***	1.905***
	(0.164)	(0.404)	(0.226)	(0.331)
N	17038	17038	7506	7506
R^2 (within)	0.107	0.112	0.083	0.092

Notes: Standard errors in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.1$

Model (2) includes the control variables, the $ARating_{it}$ and $RARating1st10_{it}$ variables, the rating dispersion variable ($ADispersion_{it}$), and its interactions with our two independent variables. The coefficients of the interaction terms are significant, indicating that the rating dispersion moderates both anchoring and swaying effects.

Table 4-6 Marginal Effects of Independent Variables.

Rating Dispersion ($ADispersion_{it}$)	Marginal Effects	
	Overall Average Rating ($ARating_{it}$)	Average Rating of the First 10 Reviews ($RARating1st10_{it}$)
Mean - 2σ	-0.405***	-0.007
	(0.063)	(0.022)
Mean - 1σ	-0.297***	-0.029**
	(0.045)	(0.014)
Mean	-0.190***	-0.051***
	(0.031)	(0.009)
Mean + 1σ	-0.082***	-0.073***
	(0.028)	(0.012)
Mean + 2σ	0.022	-0.094***
	(0.038)	(0.019)

To better interpret the moderating effects, we calculated the marginal effect of the $ARating_{it}$ variable and the $RARating1st10_{it}$ variable when the $ADispersion_{it}$ variable is fixed at different levels (e.g., mean, as well as 1 or 2 standard deviations below or above the mean value). As shown in Table 4-6, when the rating dispersion ($ADispersion_{it}$) increases, the absolute value of the coefficient of the overall average rating ($ARating_{it}$) decreases, indicating a declining marginal effect of the overall average rating in line with H2. At the same time, the absolute value of the coefficient of the average rating of the 10 reviews displayed in the first review page ($RARating1st10_{it}$) increases, indicating an increasing marginal effect of the first 10 reviews in line with H3. In fact, the marginal effect of the overall average rating disappears when the rating dispersion is extremely high (e.g., 2 standard deviations above the mean value), and the marginal effect of the first 10 reviews disappears when the

rating dispersion is extremely low (e.g., 2 standard deviations below the mean value). Taken together, the results of Model (2) indicate that the rating dispersion moderates the anchoring and swaying effects. Specifically, as the rating dispersion increases, the anchoring effect of the overall average rating becomes weaker while the swaying effect of the individual reviews is stronger, providing evidence for H2 and H3.

An alternative explanation for the results obtained from the main models is the greater variation in the average rating of the 10 reviews displayed in the first review page compared with the variation in the overall average rating. With a minimum of 150 ratings for an app to be included in the analysis, there may be less variation on overall average rating over time, subduing its effect on the dependent variable. To rule out this possibility, we did a robustness check utilizing the apps that had at least 1 unit change in the overall average rating during our study period. The right panel of Table 4-5 shows the coefficient estimates from the robustness check. The results are consistent with those from the main models.

In summary, we found additional evidence in Study 3 for the swaying effect of the most accessible reviews as proposed in H1b through data from a real-world setting. The study also provided evidence for the moderating effect of rating dispersion on both the anchoring effect proposed in H2 and the swaying effect proposed in H3.

4.6   General Discussion

Drawing on the heuristics-and-biases literature, we propose two competing hypotheses with regard to the anchoring effect of the overall average rating and the swaying effect of the most accessible individual reviews. In addition, building on the

belief confidence literature, we posit rating dispersion as a boundary condition for the anchoring and swaying effects. Utilizing two experiments and a panel dataset collected from Apple's App Store, we provide evidence that consumers' decisions are influenced more by the most accessible reviews than by the overall average rating, and that this swaying effect would be stronger and the anchoring effect would be weaker if consumers have lower confidence in their initial beliefs.

4.6.1 Theoretical Implications

Our findings have several theoretical contributions. First and foremost, we challenge the widely accepted view that aggregated rating statistics (such as the overall average rating) are the primary determinants of consumers' purchase decisions (e.g., Chevalier and Mayzlin 2006; Duan et al. 2008). Our research represents one of the first attempts to include and compare the impact of both the overall average rating and individual reviews on consumers' purchase decisions and product sales. Although prior research demonstrated that both can influence consumers' decisions (e.g., Babić Rosario et al. 2016; Chevalier and Mayzlin 2006), few studies have directly compared the relative impact of one with the other. We compare the effects of these two distinct types of information cues and find that comparing their relative impacts can yield interesting and novel theoretical insights that are not possible when they are studied separately. Specifically, our findings suggest that the most accessible reviews are more influential than the overall average rating in consumers' purchase decisions. Although aggregated rating statistics shape consumers' initial beliefs about a product, their purchase decisions are swayed easily by the most accessible reviews. Thus, earlier research might have

138

exaggerated the influence of aggregated rating statistics such as the overall average rating.

In addition, our results reveal the importance of incorporating insights from the heuristics-and-biases literature into consumer behavior research, especially in the context of online reviews. Because consumers rely on heuristics to reduce cognitive efforts and make decisions efficiently, it is an important lens for studying how consumers make sense of online reviews in their purchase decisions. Based on the anchoring heuristic, consumer's evaluations and judgments should be biased toward the anchor (Tversky and Kahneman 1974), while the availability heuristic suggests that consumers rely on instances or immediate examples that easily come to mind as the basis for their decision making (Tversky and Kahneman 1973). Applied to our context, these two well-supported heuristics from the judgement and decision-making literature offer opposite predictions with regard to the relative impact of the overall average rating that serves as the anchor and the most accessible reviews that provide concrete consumption experiences. Our findings provide evidence for the availability heuristic in general rather than the anchoring heuristic, suggesting that not all heuristics are created equal, and that some heuristics are more applicable than others are in online word-of-mouth. This study also complements existing research exploring consumer biases in decision-making (Ho et al. 2017; Yin et al. 2016).

Finally, extending the concept of belief confidence from the metacognition literature, our findings reveal that rating dispersion could be a boundary condition for both anchoring and swaying effects. Consumers form their initial beliefs based on the overall average rating of a product, and they may develop different levels of confidence

in their initial beliefs based on rating dispersions (Petrocelli et al. 2007b; Petty et al. 2007; Yin et al. 2016). Rating dispersion plays a critical role in both anchoring and swaying effects, because the likelihood of consumers' purchase decisions being biased toward the anchor depends on the strength of the anchor, and that consumers' reliance on vivid and concrete individual reviews also depends on the confidence in their initial beliefs (and subsequent attention paid to the individual reviews). Our results suggest that although consumers' purchase decisions could be swayed by the most accessible reviews, the swaying effect might vary under different conditions. In addition, the anchoring effect of the overall average rating may also arise under certain situations (e.g., when the dispersion of ratings is low).

4.6.2   *Practical Implications*

Our findings also provide practical implications for product manufacturers, retailers, and review platforms. First, the swaying effect of the most accessible reviews implies that consumers tend to place more emphasis on easily accessible reviews rather than the aggregated rating statistics when they make purchase decisions. Therefore, consumers' evaluations of products might be closer to the consumption experiences shared in the most accessible reviews. If product manufacturers and online retailers gauge consumers' interest and purchase intention based primarily on the aggregated rating statistics, this practice might be misguided. For instance, for even a product with an average rating of 4 or 4.5 stars, it is likely that consumers develop a much lower intention to purchase the product if they find the most helpful or most recent reviews to be largely negative. While the overall average rating at the product level is not easy to change because it is calculated based on all the historical ratings, the first few reviews that

consumers see are constantly changing over time. Our results suggest that product manufacturers and online retailers can better influence prospective consumers by focusing on and dealing with the first few reviews that consumers are likely to see rather than summary ratings.

Second, our investigation into a boundary condition for the swaying and anchoring effects suggests that the influence of the most accessible reviews on consumers' purchase decisions could vary when products have different rating dispersions. In addition, when a product has low dispersion in its ratings, the overall average rating of the product could overshadow the otherwise dominant swaying effect of individual reviews in consumers' purchase decisions. Thus, when product manufacturers and online retailers estimate and predict consumers' purchase decisions, they should take a more balanced view between the summary ratings and the first few reviews. Specifically, they should focus more on the first few reviews for products that have more diverse evaluations and opinions; at the same time, they should not overlook the influence of the aggregated rating statistics (such as the overall average rating) for products that have more consensus and converged evaluations from the previous users.

Third, when review platforms design the layout of product pages, they could benefit prospective consumers by giving more spotlight to individual reviews. Similar to the aggregated rating statistics such as the overall average rating, aggregating information from the most helpful or most recent reviews might be another way for review platforms to facilitate consumers' decision-making. For example, in addition to listing the most recent or helpful reviews for consumers to read, review platforms may consider

displaying the average rating of the most accessible reviews and frequently mentioned keywords from them in prominent positions.

4.6.3   Future Research

This study also presents a number of opportunities for future research. First, we assume in Study 3 that the reviews in the first review page of an app on a certain date are the most accessible reviews for that app. However, some consumers may change the display order of reviews from the default to another order. As a result, the 10 reviews on the first review page may not be the reviews that are the most accessible to a small percentage of consumers. Future research may study whether the swaying effect demonstrated in this study depends on the display order of individual reviews.

Second, we only utilized the average rating of the most accessible reviews when comparing the effect of individual reviews with that of the overall average rating. It would be interesting in future research to explore the influence of other characteristics of accessible reviews on the likelihood and extent of the swaying effect, such as the text content and keywords.

Third, we argued that the overall average rating can help consumers form initial beliefs and the distribution of ratings can shape their belief confidence. However, other information cues may also affect consumers' initial beliefs about a product (such as the brand image of a product) or influence consumers' confidence in their initial beliefs (such as the number of ratings a product receives). Future studies may want to explore these alternative information cues.

Fourth, the underlying mechanism of the swaying effect is not explored in the current research. Although our findings demonstrate that rating dispersion can be a boundary condition for both anchoring and swaying effects, the mechanism is still unknown. It would be meaningful to investigate consumers' motivation to search for and read individual reviews after their initial beliefs about a product have been formed.

4.7   Conclusion

Most online retailers and review platforms display aggregated rating statistics (such as the overall average rating) in a prominent position, assuming that consumers are heavily influenced by the product's overall average rating that is the most salient and all-encompassing signal of product quality (De Langhe et al. 2015). In addition to seeing aggregated ratings, consumers also consult individual reviews to make purchase decisions, but very few studies examined the effect of individual reviews on consumers' purchase decisions or product sales. In the current research, we take an initial step in that direction, by comparing the relative impact of overall average rating of the product with the most accessible reviews on product sales. Building on the heuristics-and-biases and belief confidence literatures, we propose two competing hypotheses and propose the rating dispersion as a boundary condition for both anchoring and swaying effects. Through two experiments and an archival analysis, we provide converging evidence for the swaying effect of the most accessible reviews, and the moderating role of rating dispersion. Our paper challenges the implicit assumption that the overall average rating is the most predominant determinant of product sales, and finds that consumers' purchase decisions are easily swayed by the most accessible reviews.

APPENDIX A. VARIABLES MEASURED IN CHAPTER 2 AND

APPENDIX B

Review helpfulness: (Sen and Lerman 2007) (used in Study 2, 3, 4A, and 4B)

Assuming that you were considering purchasing App 1 (2, or 3) in real life, how would you describe the review above?

- not at all helpful / very helpful

- not at all useful / very useful

- not at all informative / very informative

Perceived personal pronouns: (Pennebaker et al. 2003; Simmons et al. 2005) (used in manipulation check of Study 2)

In your opinion, how much does the above review contain each of the following word categories? <none / a lot>

- First-person pronouns (e.g., I, me)

- Second-person pronouns (e.g., you, your)

Review two-sidedness: (Jensen et al. 2013) (used in manipulation check of Study 2, Study 3, and the supplementary study)

Based on the review above, to what extent do you think this review is?

- very one-sided / very two-sided

- listing only pros or only cons / listing both pros and cons to a similar extent

- containing consistent opinions / containing conflicting opinions

Perceived empathic concern: (Goldstein et al. 2014; Plank et al. 1996; Toi and Batson 1982) (used in Study 3, 4A, and 4B)

Based on the reviews above, please answer the following questions: <not at all / very much>

- When reading about this review, to what extent do you think this reviewer understands your feelings?

- When reading about this review, to what extent do you think this reviewer is concerned about you?

- When reading about this review, to what extent do you think this reviewer empathizes with you?

Perceived persuasion motives: (Campbell and Kirmani 2000; Williams et al. 2004) (used in Study 3, 4A, and 4B)

Based on the review above, to what extent do you disagree or agree with the following? <strongly disagree / strongly agree>

- While I was reading the review, I thought it was pretty obvious that the reviewer was trying to influence me.

- I think this reviewer was using tricks to persuade me.

- This reviewer used inappropriate tactics in presenting information.

- While I was reading the review, I thought the reviewer was writing the review to manipulate potential users.

The following measures are used in the supplementary study (see Appendix B).

Perceived politeness: (Bargh et al. 1996; Chen and Lurie 2013)

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

- not at all polite / very polite

- not at all courteous / very courteous

- not at all respectful / very respectful

Perceived psychological closeness: (Gino and Galinsky 2012)

Based on the review above, to what extent do you disagree or agree with the following?

<strongly disagree / strongly agree>

- I would enjoy having the reviewer as a friend.

- I feel I am similar to the reviewer.

- I feel I am related to the reviewer.

- I feel I am psychologically close to the reviewer.

Perceived interpersonal closeness: (Aron et al. 1992)

Base on the review above, please circle the picture that best describes *how closely you feel* to the reviewer.

Note: Please treat:

"**Self**" as your self

"**Other**" as the reviewer who wrote the review
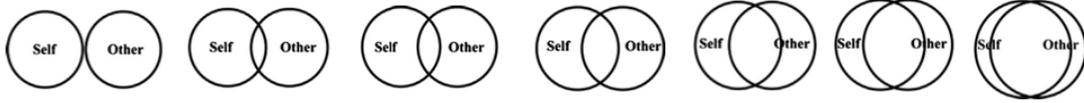
Perceived altruism: (Price et al. 1995)

In your opinion, how important is each of the following for this reviewer when he/she was writing the review? <not at all important / very important>

- To help other consumers.

- To share his/her opinions.

- To give to others.

- To be unselfish.

Review comprehensiveness: (Yang et al. 2005)

Based on the review above, to what extent do you disagree or agree with the following statements about the information provided in the review? <strongly disagree / strongly agree>

- This review offers complete descriptions about the app.

- This review offers complete content.

- This review offers sufficient information about the app.

- This review offers detailed app information.

Review valence: (MacKenzie and Lutz 1989)

Based on the review above, how would you describe the reviewer's feelings regarding the experience he/she wrote about?

- very negative / very positive

- very unfavorable / very favorable

- very unpleasant / very pleasant

APPENDIX B. SUPPLEMENTARY STUDY IN CHAPTER 2

We conducted this supplementary study to examine alternative explanations for the effect of personal pronouns and the moderating effect of review two-sidedness. In Studies 3 and 4, we measured perceived empathic concern and perceived persuasion motives as probable mechanisms underlying the main effect of personal pronouns on review helpfulness. At the same time, the positive effect of second-person (vs. first-person) pronouns may also arise from perceived politeness (i.e., the extent to which one is perceived to be respectful and considerate of other people) (Hill et al. 1986), perceived psychological closeness (i.e., perceptions of attachment and connections to others) (Gino and Galinsky 2012), perceived interpersonal closeness (i.e., one's sense of interpersonal interconnectedness and the inclusion of others in selves) (Aron et al. 1992; Berscheid et al. 1989), and perceived altruism (i.e., the motivation of helping others and increasing the welfare of the person in need) (Batson et al. 1991). In addition, our manipulation of review two-sidedness might also vary review comprehensiveness (i.e., the adequacy and completeness of information), representing a possible confound (Cheung et al. 2008; Yang et al. 2005). We measured these additional variables in this supplementary experiment to explore their possible roles.

Similar to the main studies, we manipulated personal pronouns within-subjects at two levels (first-person pronouns vs. second-person pronouns). In addition, we manipulated review two-sidedness at three levels (one-sidedness vs. low two-sidedness vs. high two-sidedness) and review valence at three levels (negative vs. neutral vs. positive). These two factors are not entirely independent as a review cannot be high in

149

two-sidedness and positive or negative in valence at the same time. In the end, we had five between-subjects conditions in total: negative one-sided, positive one-sided, negative low two-sided, positive low two-sided, and neutral high two-sided conditions.[8]

B.1 Stimulus Materials

We developed stimuli for five conditions based on treatment reviews in Studies 2 and 3. First, we kept treatment reviews used in Study 2's one-sided condition as positive one-sided condition, Study 2's two-sided condition as neutral high two-sided condition, and Study 3's low two-sided condition as positive two-sided condition in this study. Next, we constructed stimuli for negative one-sided condition and negative low two-sided condition in this study based on two sets of treatment reviews used in Study 2's two-sided condition. Similar to Study 3, for negative one-sided condition, we increased the number of negative statements from 2 to 4 (out of 4) and removed positive statements; for negative low two-sided condition, we increased the number of negative statements from 2 to 3 and decreased the number of positive statements from 2 to 1. Substantive content was held identical. All review stimuli are presented in Table B-1.

---

[8] The other four combinations do not exist: neutral one-sided, neutral low two-sided, negative high two-sided, and positive high two-sided conditions.

Table B-1 Review Stimuli in Supplementary Study.

#	Review Written in First-Person Pronouns	Review Written in Second-Person Pronouns

Negative One-Sided Condition

1 **Pros:** None.

Cons:
- This app *doesn't* help **me** focus on **my** work.
- It's *disappointing* that **I** *cannot* observe how **I** have been doing to meet goals each week.
- It *doesn't* keep the number of sessions **I** have accomplished or categorize **my** tasks.
- It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.

Pros: None.

Cons:
- This app *doesn't* help **you** focus on **your** work.
- It's *disappointing* that **you** *cannot* observe how **you** have been doing to meet goals each week.
- It *doesn't* keep the number of sessions **you** have accomplished or categorize **your** tasks.
- It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.

2 **Pros:** None.

Cons:
- The timer *fails* to help **me** say "no" to **my** incoming messages.
- It allows **me** to get **my** work done but get *no* break.
- The ticking sound is *unclear*, so **I** *cannot* remember when **I** should take a rest.
- In addition, it *cannot* sync **my** tracked data between **my** phone and computer.

Pros: None.

Cons:
- The timer *fails* to help **you** say "no" to **your** incoming messages.
- It allows **you** to get **your** work done but get *no* break.
- The ticking sound is *unclear*, so **you** *cannot* remember when **you** should take a rest.
- In addition, it *cannot* sync **your** tracked data between **your** phone and computer.

Positive One-Sided Condition

1 **Pros:**
- This app helps **me** focus on **my** work
- It's great that **I** can observe how **I** have been doing to meet goals each week.
- It keeps the number of sessions **I** have accomplished and categorizes **my** tasks.
- It's easy to organize **my** own time because **I** can customize the timer.

Cons: None.

Pros:
- This app helps **you** focus on **your** work.
- It's great that **you** can observe how **you** have been doing to meet goals each week.
- It keeps the number of sessions **you** have accomplished and categorizes **your** tasks.
- It's easy to organize **your** own time because **you** can customize the timer.

Cons: None.

2 **Pros:**
- The timer helps **me** say "no" to **my** incoming messages.
- It allows **me** to get **my** work done but still get a break.
- The ticking sound is clear, so **I** can remember when **I** should take a rest.
- In addition, it can sync **my** tracked data between **my** phone and computer.

Cons: None.

Pros:
- The timer helps **you** say "no" to **your** incoming messages.
- It allows **you** to get **your** work done but still get a break.
- The ticking sound is clear, so **you** can remember when **you** should take a rest.
- In addition, it can sync **your** tracked data between **your** phone and computer.

Cons: None.

Negative Low Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work **Cons:** • It's *disappointing* that **I** *cannot* observe how **I** have been doing to meet goals each week. • It *doesn't* keep the number of sessions **I** have accomplished or categorize **my** tasks. • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. **Cons:** • It's *disappointing* that **you** *cannot* observe how **you** have been doing to meet goals each week. • It *doesn't* keep the number of sessions **you** have accomplished or categorize **your** tasks. • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. **Cons:** • It allows **me** to get **my** work done but get *no* break. • The ticking sound is *unclear*, so **I** *cannot* remember when **I** should take a rest. • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. **Cons:** • It allows **you** to get **your** work done but get *no* break. • The ticking sound is *unclear*, so **you** *cannot* remember when **you** should take a rest. • It *cannot* sync **your** tracked data between **your** phone and computer.
Positive Low Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. • It keeps the number of sessions **I** have accomplished and categorizes **my** tasks. **Cons:** • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. • It keeps the number of sessions **you** have accomplished and categorizes **your** tasks. **Cons:** • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. • The ticking sound is clear, so **I** can remember when **I** should take a rest. **Cons:** • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. • The ticking sound is clear, so **you** can remember when **you** should take a rest. **Cons:** • It *cannot* sync **your** tracked data between **your** phone and computer.

		Pros:	Pros:
Neutral High Two-Sided Condition	1	**Pros:** • This app helps **me** focus on **my** work • It's great that **I** can observe how **I** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **I** have accomplished or categorize **my** tasks. • It's *not* easy to organize **my** own time because **I** *cannot* customize the timer.	**Pros:** • This app helps **you** focus on **your** work. • It's great that **you** can observe how **you** have been doing to meet goals each week. **Cons:** • It *doesn't* keep the number of sessions **you** have accomplished or categorize **your** tasks. • It's *not* easy to organize **your** own time because **you** *cannot* customize the timer.
	2	**Pros:** • The timer helps **me** say "no" to **my** incoming messages. • It allows **me** to get **my** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **I** *cannot* remember when **I** should take a rest. • It *cannot* sync **my** tracked data between **my** phone and computer.	**Pros:** • The timer helps **you** say "no" to **your** incoming messages. • It allows **you** to get **your** work done but still get a break. **Cons:** • The ticking sound is *unclear*, so **you** *cannot* remember when **you** should take a rest. • It *cannot* sync **your** tracked data between **your** phone and computer.

Notes: emphases and italics added for illustration purpose only (i.e., not shown to participants).

B.2 Procedure and Measures

202 respondents (120 male) from Amazon Mturk participated in this study with compensation. 99 percent were originally from the United States, 68 percent achieved bachelor's degree or above as the highest education level, and the average age was 35.

Each participant was randomly assigned into one of the five conditions. The cover story and procedure in this study were similar as in Study 2. After reading each review, we asked participants to report their perception of the reviewer's politeness using three items adapted from Chen and Lurie (2013) and Bargh et al. (1996), their perception of the psychological closeness with the reviewer using four items adapted from Gino and Galinsky (2012), their perception of the interpersonal closeness with the reviewer using

an item adapted from Aron et al. (1992), their perception of the reviewer's altruism using

four items adapted from Price et al. (1995). In addition, we asked participants to evaluate

review comprehensiveness using three items adapted from Yang et al. (2005). As a

manipulation check, participants also rated the level of review two-sidedness using the

same items as in Study 2, and the review valence using three items adapted from

MacKenzie and Lutz (1989). See Appendix A for all the measures.

B.3   Results

We first conducted a manipulation check of review two-sidedness and review

valence. ANCOVA results revealed that perceived review two-sidedness in the one-sided,

low two-sided, and high two-sided conditions followed the expected pattern ($M = 4.12$

vs. 5.35 vs. 6.08, $F(1, 196) = 18.14$, $p < 0.001$). In addition, perceived review valence in

the negative, neutral, and positive conditions also followed the expected pattern ($M =$

4.74 vs. 6.09 vs. 7.22, $F(1, 196) = 83.63$, $p < 0.001$). Therefore, our manipulations of

both variables were successful.

We then used ANCOVA to examine the effect of personal pronouns on perceived

politeness, perceived psychological closeness, and perceived interpersonal closeness. In

the series of ANCOVA analyses, we entered each of these variables as the dependent

variable, personal pronouns as a within-subjects factor, the level of review two-sidedness

and the level of review valence as between-subjects factors, and treatment order as a

covariate. Results revealed that the effect of personal pronouns on perceived politeness,

perceived psychological closeness, and perceived interpersonal closeness did not reach

significance (politeness: $M = 6.88$ vs. 6.90, $F(1, 196) = 0.03$, $p = 0.870$; psychological

closeness: $M = 5.39$ vs. 5.43, $F(1, 196) = 0.59$, $p = 0.442$; interpersonal closeness: $M = 3.52$ vs. 3.49, $F(1, 196) = 0.22$, $p = 0.640$). Thus, these results suggest that perceived politeness, perceived psychological closeness, and perceived interpersonal closeness are not likely to mediate the effect of personal pronouns on review helpfulness.

On the other hand, results revealed that a reviewer's use of second-person (vs. first-person) pronouns had a positive effect on perceived altruism ($M = 6.35$ vs. 6.54, $F(1, 196) = 6.67$, $p = 0.011$). As a result, we cannot rule out perceived altruism as a probable underlying mechanism. However, this alternative explanation does not conflict with our arguments on the positive relationship between the use of second-person (vs. first-person) pronouns and perceived empathic concern because perceived altruism might be a natural consequence of perceived empathic concern. For example, previous literature found that empathic feelings can boost one's motivation to help others and increase others' welfare (e.g., Batson et al. 1991; Batson et al. 1997).

In addition, we conducted a similar ANCOVA analysis to explore the effect of manipulated review two-sidedness on review comprehensiveness. Results showed that perceived review comprehensiveness did not differ significantly across the one-sided, low two-sided, and high two-sided conditions ($M = 6.00$ vs. 6.02 vs. 6.00, $F(1, 196) = 0.027$, $p = 0.948$). Hence, review two-sidedness is not confounded with review comprehensiveness.

B.4 Discussion

This supplementary study explored a few additional mechanisms that might underlie the main effect of personal pronouns and the moderating effect of review two-

sidedness. We found that personal pronouns in the review did not vary readers' perceptions of reviewer politeness, as well as their psychological closeness and interpersonal closeness with the reviewer. On the other hand, results revealed that the effect of personal pronouns might be explained by readers' perception of reviewer altruism. However, this is in line with our empathy-based arguments, such that the more second-person pronouns reviewers use in reviews, the more likely that they are perceived to be empathic and concerned about readers (and correspondingly more willing to help readers). In addition, our results indicated that the manipulation of review two-sidedness did not influence review comprehensiveness. Taken together, this supplementary study indicated that perceived politeness, perceived psychological closeness, perceived interpersonal closeness, and review comprehensiveness are not likely to account for the effects observed in the main experiments, and that an explanation driven by perceived altruism is consistent with the positive mediating process (i.e., perceived empathic concern) discussed in the paper.

APPENDIX C. PRETEST OF REVIEW TITLES IN CHAPTER 3

Note: all Likert and sematic differential items were measured using 9-point scales, unless specified otherwise.

Instructions

Please read the following description carefully:

In this task, we would like to get your help in evaluating products from Amazon.com.

Amazon.com allows users to browse and purchase products online. The website sells over 480 million products in the USA under dozens of departments including Sports, Electronics, Books, etc. Amazon.com enables users to rate any product along a scale of 1 to 5 stars. Users can also submit a text review to describe their experience with the product that they have purchased and used.

Continued on the next screen.

Instructions – Continued

Imagine you are planning to purchase a particular type of wireless mouse from Amazon.com. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. You would prefer a mouse that is not only compact but also foldable, so that you can easily take it with you anywhere you go.

You open Amazon.com, type in "compact foldable wireless mouse" and click the "Search" button.

Continued on the next screen.

Instructions – Continued

Your search returns a compact foldable wireless mouse, which costs $23.99.

At this point, you don't know which one to purchase. Therefore, you decide to read some user reviews.

Continued on the next screen.

Instructions – Continued

Imagine that you don't have enough time to read all of the reviews, but you can read titles of some reviews first.

Below are titles of the 12 most recent reviews of the wireless mouse.

<12 review titles>

// *Extremity:* (Lee et al. 2009)
In your opinion, how negative are these titles?
- not at all very negative / very negative

In your opinion, how positive are these titles?
- not at all positive / very positive

// 12 review titles:

Positive titles: great product, fabulous, joyful experience, attractive product, terrific, and wise choice.

Negative titles: it's worthless, disturbing, depressing purchase, useless one, undesirable, and terrible product.

APPENDIX D. PRETEST OF REVIEWS IN CHAPTER 3

Note: all Likert and sematic differential items were measured using 9-point scales, unless specified otherwise.

Instructions

Please read the following description carefully:

In this task, we would like to get your help in evaluating products from Amazon.com.

Amazon.com allows users to browse and purchase products online. The website sells over 480 million products in the USA under dozens of departments including Sports, Electronics, Books, etc. Amazon.com enables users to rate any product along a scale of 1 to 5 stars. Users can also submit a text review to describe their experience with the product that they have purchased and used.

Continued on the next screen.

Instructions – Continued

Imagine you are planning to purchase a particular type of wireless mouse from Amazon.com. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. You would prefer a mouse that is not only compact but also foldable, so that you can easily take it with you anywhere you go.

You open Amazon.com, type in "compact foldable wireless mouse" and click the "Search" button.

Continued on the next screen.

Instructions – Continued

Your search returns a compact foldable wireless mouse, which costs $23.99.

At this point, you don't know which one to purchase. Therefore, you decide to

read some user reviews.

Continued on the next screen.

Instructions – Continued

Imagine that you don't have enough time to read all of the reviews, but you have time to read the 6 most recent reviews of the wireless mouse.

NOTE: You will be reading and evaluating the text reviews one at a time. Please make sure to read the entire review carefully before answering any questions.

The task starts on the next screen.

// Each participant is randomly assigned to read one version (positive or negative) of the reviews in each set.

<Review 1>

// *Extremity:* (Lee et al. 2009)
In your opinion, how negative is this review above?
- not at all very negative / very negative

In your opinion, how positive is this review above?
- not at all positive / very positive

// *Information quantity:* (Gao et al. 2012)
In your opinion, how much information was presented in this review above?
- very little information / a great deal of information
- very few details / very many details

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

// *Quality:* (McKinney et al. 2002)
- very poor quality / very good quality
- very poor content / very good content
- very incomplete / very complete

// *Reading difficulty:* (Ermakova et al. 2014; Hall and Hanna 2004)
- very hard to read / very easy to read
- very hard to understand / very easy to understand

> // *Appropriateness:* (Glikson et al. 2017)
> - written inappropriately / written appropriately
> - very poorly articulated / very well articulated
>
> // *Realism:* (Mafael et al. 2016)
> - not at all realistic / very realistic
> - not at all real / very real

Repeat for the other 5 reviews.

// Content of reviews in the 6 sets.

Set #	Positive Version	Negative Version
1	This is a great mouse and it works well. The mouse has the curved left side for the thumb, so it's very comfortable. Moreover, it allows me to change how quickly the cursor moves across my screen.	This is a worthless mouse and it doesn't work well. The mouse doesn't have the curved left side for the thumb, so it's very uncomfortable. Moreover, it doesn't allow me to change how quickly the cursor moves across my screen.
2	Very good wireless mouse. I like the side buttons, which are programmed to go back or forward on web browsers by default. The mouse has a setup software, so there is an easy way to reprogram the buttons.	Very bad wireless mouse. I don't like the side buttons, which are programmed to go back or forward on web browsers by default. The mouse has no setup software, so there is no easy way to reprogram the buttons.
3	It's easy to use. I purchased this item a few months ago and I am pleased with its performance. The tracking on this mouse is good. It's a desirable mouse for the price. I would definitely recommend it.	It's difficult to use. I purchased this item a few months ago and I am not pleased with its performance. The tracking on this mouse is poor. It's an undesirable mouse for the price. I would definitely not recommend it.
4	High quality. It is comfortable to use, especially if it's being used for over an hour in one sitting. Also, it is durable as the mouse was knocked off my desk and shown no clear sign of damage.	Poor quality. It isn't comfortable to use, especially if it's being used for over an hour in one sitting. Also, it isn't durable as the mouse was knocked off my desk and shown a clear sign of damage.
5	The mouse functions well. One feature that I found useful for saving battery life is the mouse turns off automatically after a long time of non-use. It is convenient for someone who walks away from their computer often.	The mouse functions poorly. One feature that I found harmful for saving battery life is the mouse doesn't turn off automatically after a long time of non-use. It isn't convenient for someone who walks away from their computer often.
6	Good value for the price. It includes a battery with the product, so you can use it immediately. It connects to my laptop very quickly. And it is responsive without any lag when I move it.	Poor value for the price. It doesn't include a battery with the product, so you cannot use it immediately. It connects to my laptop very slowly. And it isn't responsive with lags when I move it.

APPENDIX E. STUDY 1 IN CHAPTER 3

Note: all Likert and sematic differential items were measured using 9-point scales, unless specified otherwise.

Instructions

Please read the following description carefully:

In this task, we would like to get your help in evaluating products from Amazon.com.

Amazon.com allows users to browse and purchase products online. The website sells over 480 million products in the USA under dozens of departments including Sports, Electronics, Books, etc. Amazon.com enables users to rate any product along a scale of 1 to 5 stars. Users can also submit a text review to describe their experience with the product that they have purchased and used.

Continued on the next screen.

Instructions – Continued

Imagine you are planning to purchase a particular type of wireless mouse from Amazon.com. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. You would prefer a mouse that is not only compact but also foldable, so that you can easily take it with you anywhere you go.

You open Amazon.com, type in "compact foldable wireless mouse" and click the "Search" button.

Continued on the next screen.

Instructions – Continued

Your search returns **2 different compact foldable wireless mice**, each costing **$23.99**. We will refer to these 2 mice by letters: **Mouse A** and **Mouse B**.

At this point, you don't know which one to purchase. Therefore, you decide to consult their rating profiles – summary statistics of user ratings, and perhaps read some user reviews.

Continued on the next screen.

Instructions – Continued

For **each mouse**, you will be provided with its rating profile – summary statistics of ratings from previous users who have purchased it. After seeing the rating profiles of two wireless mice, you will be asked to give your impressions of the mice.

In total, you will see 2 mice. The task starts on the next screen.

| **Mouse A** | **Mouse B** |
| <Rating Profiles> | <Rating Profiles> |

// Initial beliefs: (Darke and Ritchie 2007)
What is your overall opinion of **Mouse A** based on its rating profile on the top left of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

| **Mouse A** | **Mouse B** |
| <Rating Profiles> | <Rating Profiles> |

// Initial beliefs:
What is your overall opinion of **Mouse B** based on its rating profile on the top right of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

Mouse A
<Rating Profiles>

Mouse B
<Rating Profiles>

// Choose the superior product:
Assume you are in a hurry and only have time to read reviews from one of the two mice. Based on their rating profiles above, which mouse would you choose to find out more information about it?
- definitely choose Mouse A / definitely choose Mouse B

You have chosen Mouse A/B. See its rating profile below. We will ask you some questions about Mouse A/B.

Mouse A/B
<Rating Profiles>

Amazon.com displays the 6 most recent reviews of Mouse A/B on the first review page. **Imagine that you don't have enough time to read all of the reviews, and you only have time to read 3 of them.**

Below are the titles of the 6 most recent reviews of Mouse A/B. Please **pick 3 reviews that you are most interested in reading**.

<6 review titles >

Mouse A/B
<Rating Profiles>

Below are the **3 reviews** of Mouse A/B that you have picked. Please read them carefully before answering any questions.

< 3 reviews>

From this screen, you will be asked to provide your evaluation of the individual reviews of Mouse A/B.

<div align="center"><Review 1></div>

// *Review helpfulness:* (Chen and Lurie 2013; Sen and Lerman 2007)
Assuming that you were thinking about purchasing Mouse A/B in real life, how likely would you be to use this review in your decision-making?
- very unlikely / very likely

How much influence would this review have on your decision?
- very little influence / a great deal of influence

Repeat for the other 2 reviews.

APPENDIX F. STUDY 2 IN CHAPTER 3

Note: all Likert and sematic differential items were measured using 9-point scales, unless specified otherwise.

Instructions

Please read the following description carefully:

In this task, we would like to get your help in evaluating products from Amazon.com.

Amazon.com allows users to browse and purchase products online. The website sells over 480 million products in the USA under dozens of departments including Sports, Electronics, Books, etc. Amazon.com enables users to rate any product along a scale of 1 to 5 stars. Users can also submit a text review to describe their experience with the product that they have purchased and used.

Continued on the next screen.

Instructions – Continued

Imagine you are planning to purchase a particular type of wireless mouse from Amazon.com. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. You would prefer a mouse that is not only compact but also foldable, so that you can easily take it with you anywhere you go.

You open Amazon.com, type in "compact foldable wireless mouse" and click the "Search" button.

Continued on the next screen.

<div align="center">

Mouse A   **Mouse B**
<Rating Profiles>   <Rating Profiles>

</div>

// Initial beliefs: (Darke and Ritchie 2007)
What is your overall opinion of **Mouse A** based on its rating profile on the top left of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

Mouse A	**Mouse B**
<Rating Profiles>	<Rating Profiles>

// Initial beliefs:

What is your overall opinion of **Mouse B** based on its rating profile on the top right of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

We have randomly selected one of the two mice – Mouse A/B (the one on the left/right) – for you to evaluate next.

// Rating profiles of the randomly selected mouse are highlighted.

Mouse A	**Mouse B**
<Rating Profiles>	<Rating Profiles>

Mouse A/B
<Rating Profiles>

Amazon.com displays the 6 most recent reviews of Mouse A/B on the first review page. **Imagine that you don't have enough time to read all of the reviews, and you only have time to read 3 of them.**

Below are the titles of the 6 most recent reviews of Mouse A/B. Please **pick 3 reviews that you are most interested in reading**.

<6 review titles >

Mouse A/B
<Rating Profiles>

Below are the **3 reviews** of Mouse A/B that you have picked. Please read them carefully before answering any questions.

< 3 reviews>

From this screen, you will be asked to provide your evaluation of the individual reviews of Mouse A/B.

<Review 1>

// *Review helpfulness:* (Chen and Lurie 2013; Sen and Lerman 2007)
Assuming that you were thinking about purchasing Mouse A/B in real life, how likely would you be to use this review in your decision-making?
- very unlikely / very likely

How much influence would this review have on your decision?
- very little influence / a great deal of influence

Repeat for the other 2 reviews.

APPENDIX G. STUDY 3 IN CHAPTER 3

Note: all Likert and sematic differential items were measured using 9-point scales, unless specified otherwise.

Instructions

Please read the following description carefully:

In this task, we would like to get your help in evaluating products from Amazon.com.

Amazon.com allows users to browse and purchase products online. The website sells over 480 million products in the USA under dozens of departments including Sports, Electronics, Books, etc. Amazon.com enables users to rate any product along a scale of 1 to 5 stars. Users can also submit a text review to describe their experience with the product that they have purchased and used.

Continued on the next screen.

Instructions – Continued

Imagine you are planning to purchase a particular type of wireless mouse from Amazon.com. A wireless mouse is a computer mouse that needs no wires to send signals from the mouse to a computer. You would prefer a mouse that is not only compact but also foldable, so that you can easily take it with you anywhere you go.

You open Amazon.com, type in "compact foldable wireless mouse" and click the "Search" button.

Continued on the next screen.

Instructions – Continued

Your search returns **2 different compact foldable wireless mice**, each costing **$23.99**. We will refer to these 2 mice by letters: **Mouse A** and **Mouse B**.

At this point, you don't know which one to purchase. Therefore, you decide to consult their rating profiles – summary statistics of user ratings, and perhaps read some user reviews.

Continued on the next screen.

Instructions – Continued

For **each mouse**, you will be provided with its rating profile – summary statistics of ratings from previous users who have purchased it. After seeing the rating profiles of two wireless mice, you will be asked to give your impressions of the mice.

In total, you will see 2 mice. The task starts on the next screen.

Mouse A	**Mouse B**
<Rating Profiles>	<Rating Profiles>

// *Initial beliefs:* (Darke and Ritchie 2007)
What is your overall opinion of **Mouse A** based on its rating profile on the top left of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

Mouse A	**Mouse B**
<Rating Profiles>	<Rating Profiles>

// *Initial beliefs:*
What is your overall opinion of **Mouse B** based on its rating profile on the top right of this page?
- very bad / very good
- very positive / very negative
- very unfavorable / very favorable

> **Mouse A**
> <Rating Profiles>
>
> **Mouse B**
> <Rating Profiles>
>
>   Assume **Mouse A** appears first in your search result and catches your attention first, promoting you to check out **Mouse A's** reviews first.

> **Mouse A**
> <Rating Profiles>
>
>   Above is the rating profile of Mouse A. Amazon.com displays the 6 most recent reviews of Mouse A on the first review page.
>
>   Imagine that you don't have enough time to read all of the reviews, and you only have time to read 3 of them.
>
>   Below are the titles of the 6 most recent reviews of Mouse A. Please **pick 3 reviews that you are most interested in reading.**
>
> <6 review titles >

> **Mouse A**
> <Rating Profiles>
>
>   Below are the **3 reviews** of Mouse A that you have picked. Please read them carefully before answering any questions.
>
> < 3 reviews>

>   From this screen, you will be asked to provide your evaluation of the individual reviews of Mouse A.
>
> <Review 1>
>
> // *Review helpfulness:* (Chen and Lurie 2013; Sen and Lerman 2007)
>   Assuming that you were thinking about purchasing Mouse A in real life, how likely would you be to use this review in your decision-making?
> - very unlikely / very likely

How much influence would this review have on your decision?
- very little influence / a great deal of influence

Repeat for the other 2 reviews.

APPENDIX H. VARIABLES MEASURED IN THE PRETEST OF

CHAPTER 4

Assume that you were considering purchasing the camera. Using the scales below, how would you describe the review above?

Extremity: (Lee et al. 2009)

- not at all positive / very positive

- not at all pleasant / very pleasant

(or)

- not at all negative / very negative

- not at all unpleasant / very unpleasant

Information quantity: (Gao et al. 2012)

- contains very little information / contains a great deal of information

- information contained in the review was not thorough at all / information contained in the review was very thorough

Concreteness: (Keller and Block 1997)

- not at all concrete / very concrete

- not at all specific / very specific

Helpfulness: (Sen and Lerman 2007)

- not at all helpful / very helpful

- not at all informative / very informative

Emotional intensity: (Jensen et al. 2013)

- contains little emotion / contains a great deal of emotion

- contains no feelings / contains a lot of feelings

Realism: (Mafael et al. 2016)

- not at all realistic / very realistic

- not at all real / very real

Reading difficulty: (Ermakova et al. 2014; Hall and Hanna 2004)

- very hard to read / very easy to read

- very hard to understand / very easy to understand

APPENDIX I. VARIABLES MEASURED IN STUDY 1 AND

STUDY 2 OF CHAPTER 4

Purchase intention: (Dodds et al. 1991; Goldberg and Gorn 1987)

Based on the information of Camera Model A/B, please answer the following questions.

- If you were thinking of buying a digital camera, how likely is it that you would buy Camera Model A/B?

- How likely is it that you would consider purchasing Camera Model A/B?

- How likely is it that Camera Model A/B would be a good choice for you?

Choice between two products:

Given a choice between the two cameras, which camera would you choose? (8-point scale)

- definitely choose Camera Model A / definitely choose Camera Model B

Average rating:

Study 1:

Can you recall the average rating of Camera Model A/B based on hundreds of its prior customer reviews? (5-point scale)

- 3 stars / 3.5 stars / 4 stars / 4.5 stars / 5 stars

Study 2:

Can you recall the average rating of Camera Model A/B based on hundreds of its prior customer reviews? (5-point scale)

- 2 stars / 2.5 stars / 3 stars / 3.5 stars / 4 stars

Review valence: (MacKenzie and Lutz 1989)

Using the scales below, overall, how would you describe the above 3 reviews as a whole?

- expresses very bad feelings about the camera / expresses very good feelings about the camera

- expresses very unfavorable feelings about the camera / expresses very favorable feelings about the camera

- expresses very unpleasant feelings about the camera / expresses very pleasant feelings about the camera

Rating dispersion: (He and Bond 2015)

Based on the information of Camera Model A/B, please describe the rating distribution of Camera Model A/B.

- very close together / very spread apart

- very low dispersion / very high dispersion

REFERENCES

Angrist, J. D., and Pischke, J.-S. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Ariely, D., Loewenstein, G., and Prelec, D. 2003. ""Coherent Arbitrariness": Stable Demand Curves without Stable Preferences," *The Quarterly Journal of Economics* (118:1), pp. 73-106.

Aron, A., Aron, E. N., and Smollan, D. 1992. "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness," *Journal of Personality and Social Psychology* (63:4), p. 596.

Arriaga, X. B., and Rusbult, C. E. 1998. "Standing in My Partner's Shoes: Partner Perspective Taking and Reactions to Accommodative Dilemmas," *Personality and Social Psychology Bulletin* (24:9), pp. 927-948.

Babić Rosario, A., Sotgiu, F., De Valck, K., and Bijmolt, T. H. 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), pp. 297-318.

Baek, H., Ahn, J., and Choi, Y. 2012. "Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues," *International Journal of Electronic Commerce* (17:2), pp. 99-126.

Baek, H., Lee, S., Oh, S., and Ahn, J. 2015. "Normative Social Influence and Online Review Helpfulness: Polynomial Modeling and Response Surface Analysis," *Journal of Electronic Commerce Research* (16:4), p. 290.

Bargh, J. A., Chen, M., and Burrows, L. 1996. "Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action," *Journal of Personality and Social Psychology* (71:2), p. 230.

Basuroy, S., Chatterjee, S., and Ravid, S. A. 2003. "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," *Journal of marketing* (67:4), pp. 103-117.

Batson, C. D. 1987. "Prosocial Motivation: Is It Ever Truly Altruistic?," *Advances in Experimental Social Psychology* (20), pp. 65-122.

Batson, C. D. 1991. *The Altruism Question: Toward a Social-Psychological Answer*, Hillsdale, NJ: Erlbaum.

Batson, C. D. 1998. "Altruism and Prosocial Behavior," in *The Handbook of Social Psychology,* S.T. Fiske, D.T. Gilbert and G. Lindzey (eds.), Boston: McGraw-Hill, pp. 282-316.

Batson, C. D. 2009. "These Things Called Empathy: Eight Related but Distinct Phenomona." Cambridge, Mass.: MIT Press, pp. 3-15.

Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., and Todd, R. M. 1991. "Empathic Joy and the Empathy-Altruism Hypothesis," *Journal of Personality and Social Psychology* (61:3), p. 413.

Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., and Dawson, K. 1997. "Is Empathy-Induced Helping Due to Self-Other Merging?," *Journal of Personality and Social Psychology* (73:3), p. 495.

Batson, C. D., Sympson, S. C., Hindman, J. L., Decruz, P., Todd, R. M., Weeks, J. L., Jennings, G., and Burns, C. T. 1996. ""I've Been There, Too": Effect on Empathy of Prior Experience with a Need," *Personality and Social Psychology Bulletin* (22:5), pp. 474-482.

Bauer, R. 1967. "Consumer Behavior as Risk Taking," in *Risk Taking and Information Handling in Consumer Behavior,* D.F. Cox (ed.), Boston: Graduate School of Business Administration, Harvard University, pp. 23-33.

Baum, C. F. 2008. "Stata Tip 63: Modeling Proportions," *Stata Journal* (8:2), pp. 299-303.

Beauvois, J. L., and Joule, R. V. 1996. *A Radical Dissonance Theory*, Philadelphia, PA, US: Taylor & Francis.

Berlyne, D. E. 1970. "Novelty, Complexity, and Hedonic Value," *Perception & Psychophysics* (8:5), pp. 279-286.

Berscheid, E., Snyder, M., and Omoto, A. M. 1989. "The Relationship Closeness Inventory: Assessing the Closeness of Interpersonal Relationships," *Journal of Personality and Social Psychology* (57:5), p. 792.

Brannon, L. A., Tagler, M. J., and Eagly, A. H. 2007. "The Moderating Role of Attitude Strength in Selective Exposure to Information," *Journal of Experimental Social Psychology* (43:4), pp. 611-617.

Brechan, I. 2002. "Selective Exposure and Selective Attention: The Moderating Effect of Confidence in Attitudes and the Knowledge Basis for These Attitudes," in: *Unpublished master's thesis*. University of Florida.

Brehm, S. S., and Brehm, J. W. 2013. *Psychological Reactance: A Theory of Freedom and Control*, New York: Academic Press.

Burgoon, J. K., Guerrero, L. K., and Floyd, K. 2016. *Nonverbal Communication*, Routledge.

Campbell, M. C., and Kirmani, A. 2000. "Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent," *journal of Consumer Research* (27:1), pp. 69-83.

Campbell, R. S., and Pennebaker, J. W. 2003. "The Secret Life of Pronouns: Flexibility in Writing Style and Physical Health," *Psychological Science* (14:1), pp. 60-65.

Cao, Q., Duan, W., and Gan, Q. 2011. "Exploring Determinants of Voting for the "Helpfulness" of Online User Reviews: A Text Mining Approach," *Decision Support Systems* (50:2), pp. 511-521.

Carlson, K. A., and Russo, J. E. 2001. "Biased Interpretation of Evidence by Mock Jurors," *Journal of Experimental Psychology: Applied* (7:2), p. 91.

Case, D. O., Andrews, J. E., Johnson, J. D., and Allard, S. L. 2005. "Avoiding Versus Seeking: The Relationship of Information Seeking to Avoidance, Blunting, Coping, Dissonance, and Related Concepts," *Journal of the Medical Library Association* (93:3), p. 353.

Chaffee, S. H., Saphir, M. N., Graf, J., Sandvig, C., and Hahn, K. S. 2001. "Attention to Counter-Attitudinal Messages in a State Election Campaign," *Political Communication* (18:3), pp. 247-272.

Chaiken, S. 1980. "Heuristic Versus Systematic Information-Processing and the Use of Source Versus Message Cues in Persuasion," *Journal of Personality and Social Psychology* (39:5), pp. 752-766.

Chaiken, S., Liberman, A., and Eagly, A. H. 1989. "Heuristic and Systematic Information Processing within and Beyound the Persuasion Context," in *Unintended Thought,* J.S. Uleman and B.J. A. (eds.), New York, NY: Guilford Press, pp. 212-252.

Chaiken, S., Wood, W., and Eagly, A. H. 1996. "Principles of Persuasion," in *Social Psychology: Hand-Book of Basic Principles* E.T. Higgins and A.W. Kruglanski (eds.), New York: Guilford Press, pp. 361-399.

Chapman, G. B., and Johnson, E. J. 1999. "Anchoring, Activation, and the Construction of Values," *Organizational Behavior and Human Decision Processes* (79:2), pp. 115-153.

Chen, Z., and Lurie, N. H. 2013. "Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth," *Journal of Marketing Research* (50:4), pp. 463-476.

Cheung, C. M., Lee, M. K., and Rabjohn, N. 2008. "The Impact of Electronic Word-of-Mouth: The Adoption of Online Opinions in Online Customer Communities," *Internet Research* (18:3), pp. 229-247.

Cheung, C. M.-Y., Sia, C.-L., and Kuan, K. K. 2012. "Is This Review Believable? A Study of Factors Affecting the Credibility of Online Consumer Reviews from an Elm Perspective," *Journal of the Association for Information Systems* (13:8), p. 618.

Cheung, M. Y., Luo, C., Sia, C. L., and 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), pp. 9-38.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of marketing research* (43:3), pp. 345-354.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345-354.

Chung, C., and Pennebaker, J. W. 2007. "The Psychological Functions of Function Words," *Social Communication* (1), pp. 343-359.

Clark, R. A., and Delia, J. G. 1976. "The Development of Functional Persuasive Skills in Childhood and Early Adolescence," *Child Development* (47:4), pp. 1008-1014.

Clee, M. A., and Wicklund, R. A. 1980. "Consumer Behavior and Psychological Reactance," *Journal of Consumer Research* (6:4), pp. 389-405.

Coke, J. S., Batson, C. D., and McDavis, K. 1978. "Empathic Mediation of Helping: A Two-Stage Model," *Journal of Personality and Social Psychology* (36:7), pp. 752-766.

Corver, N., and van Riemsdijk, H. C. 2001. *Semi-Lexical Categories: The Function of Content Words and the Content of Function Words*, Walter de Gruyter.

Crowley, A. E., and Hoyer, W. D. 1994. "An Integrative Framework for Understanding Two-Sided Persuasion," *Journal of Consumer Research* (20:4), pp. 561-574.

Darke, P. R., and Ritchie, R. J. 2007. "The Defensive Consumer: Advertising Deception, Defensive Processing, and Distrust," *Journal of Marketing Research* (44:1), pp. 114-127.

David, P. 1996. "Role of Imagery in Recall of Deviant News," *Journalism & Mass Communication Quarterly* (73:4), pp. 804-820.

Davis, M. H. 1983. "Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach," *Journal of Personality and Social Psychology* (44:1), pp. 113-126.

De Langhe, B., Fernbach, P. M., and Lichtenstein, D. R. 2015. "Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings," *Journal of Consumer Research* (42:6), pp. 817-833.

Dellarocas, C. 2003. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* (49:10), pp. 1407-1424.

Ditto, P. H., and Lopez, D. F. 1992. "Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions," *Journal of Personality and Social Psychology* (63:4), pp. 568-584.

Dodds, W. B., Monroe, K. B., and Grewal, D. 1991. "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," *Journal of Marketing Research* (238:3), pp. 307-319.

Duan, W. J., Gu, G., and Whinston, A. B. 2008. "Do Online Reviews Matter? - an Empirical Investigation of Panel Data," *Decision Support Systems* (45:4), pp. 1007-1016.

Duan, W., Gu, B., and Whinston, A. B. 2008. "The Dynamics of Online Word-of-Mouth and Product Sales—an Empirical Investigation of the Movie Industry," *Journal of Retailing* (84:2), pp. 233-242.

Eagly, A. H., Chen, S., Chaiken, S., and Shaw-Barnes, K. 1999. "The Impact of Attitudes on Memory: An Affair to Remember," *Psychological Bulletin* (125:1), pp. 64-89.

Edwards, K., and Smith, E. E. 1996. "A Disconfirmation Bias in the Evaluation of Arguments," *Journal of Personality and Social Psychology* (71:1), pp. 5-24.

Elliot, A. J. 1999. "Approach and Avoidance Motivation and Achievement Goals," *Educational Psychologist* (34:3), pp. 169-189.

Erdelyi, M. H. 1974. "A New Look at the New Look: Perceptual Defense and Vigilance," *Psychological review* (81:1), p. 1.

Ermakova, T., Baumann, A., Fabian, B., and Krasnova, H. 2014. "Privacy Policies and Users' Trust: Does Readability Matter?," *Proceedings of the 20th Americas Conference on Information Systems*, pp. 1-12.

Escalas, J. E., and Stern, B. B. 2003. "Sympathy and Empathy: Emotional Responses to Advertising Dramas," *Journal of Consumer Research* (29:4), pp. 566-578.

Festinger, L. 1957. *A Theory of Cognitive Dissonance*, Stanford University Press: Stanford University Press.

Fischer, P., and Greitemeyer, T. 2010. "A New Look at Selective-Exposure Effects," *Current Directions in Psychological Science* (19:6), pp. 384-389.

Fischer, P., Greitemeyer, T., and Frey, D. 2008a. "Self-Regulation and Selective Exposure: The Impact of Depleted Self-Regulation Resources on Confirmatory Information Processing," *Journal of Personality and Social Psychology* (94:3), pp. 382-395.

Fischer, P., Jonas, E., Frey, D., and Schulz-Hardt, S. 2005. "Selective Exposure to Information: The Impact of Information Limits," *European Journal of Social Psychology* (35:4), pp. 469-492.

Fischer, P., Kastenmüller, A., Greitemeyer, T., Fischer, J., Frey, D., and Crelley, D. 2011. "Threat and Selective Exposure: The Moderating Role of Threat and Decision Context on Confirmatory Information Search after Decisions," *Journal of Experimental Psychology: General* (140:1), p. 51.

Fischer, P., Schulz-Hardt, S., and Frey, D. 2008b. "Selective Exposure and Information Quantity: How Different Information Quantities Moderate Decision Makers' Preference for Consistent and Inconsistent Information," *Journal of Ppersonality and Social Psychology* (94:2), p. 231.

Fiske, S. T., and Taylor, S. E. 1991. *Social Cognition*, Mcgraw-Hill Book Company.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., and Freling, T. 2014. "How Online Product Reviews Affect Retail Sales: A Meta-Analysis," *Journal of Retailing* (90:2), pp. 217-232.

Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp. 291-313.

Fraley, B., and Aron, A. 2004. "The Effect of a Shared Humorous Experience on Closeness in Initial Encounters," *Personal Relationships* (11:1), pp. 61-78.

Frey, D. 1986. "Recent Research on Selective Exposure to Information," in *Advances in Experimental Social Psychology,* L. Berkowitz (ed.), New York: Academic Press, pp. 41-80.

Friestad, M., and Wright, P. 1994. "The Persuasion Knowledge Model: How People Cope with Persuasion Attempts," *Journal of Consumer Research* (21:1), pp. 1-31.

Furnham, A., and Boo, H. C. 2011. "A Literature Review of the Anchoring Effect," *The Journal of Socio-Economics* (40:1), pp. 35-42.

Galinsky, A. D., and Moskowitz, G. B. 2000. "Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and in-Group Favoritism," *Journal of Personality and Social Psychology* (78:4), pp. 708-724.

Galinsky, A. D., Maddux, W. W., Gilin, D., and White, J. B. 2008. "Why It Pays to Get inside the Head of Your Opponent the Differential Effects of Perspective Taking and Empathy in Negotiations," *Psychological Science* (19:4), pp. 378-384.

Gao, J., Zhang, C., Wang, K., and Ba, S. 2012. "Understanding Online Purchase Decision Making: The Effects of Unconscious Thought, Information Quality, and Information Quantity," *Decision Support Systems* (53:4), pp. 772-781.

Gigerenzer, G., and Todd, P. M. 1999. "Fast and Frugal Heuristics: The Adaptive Toolbox," in *Simple Heuristics That Make Us Smart*, Oxford University Press, pp. 3-34.

Gilovich, T., Kerr, M., and Medvec, V. H. 1993. "Effect of Temporal Perspective on Subjective Confidence," *Journal of Personality and Social Psychology* (64:4), pp. 552-560.

Gino, F., and Galinsky, A. D. 2012. "Vicarious Dishonesty: When Psychological Closeness Creates Distance from One's Moral Compass," *Organizational Behavior and Human Decision Processes* (119:1), pp. 15-26.

Glikson, E., Cheshin, A., and Van Kleef, G. A. 2018. "The Dark Side of a Smiley: Effects of Smiling Emoticons on Virtual First Impressions," in: *Social Psychological and Personality Science* (9:5), pp. 614-625.

Goldberg, M. E., and Gorn, G. J. 1987. "Happy and Sad Tv Programs: How They Affect Reactions to Commercials," *Journal of Consumer Research* (14:3), pp. 387-403.

Goldstein, N. J., Vezich, I. S., and Shapiro, J. R. 2014. "Perceived Perspective Taking: When Others Walk in Our Shoes," *Journal of Personality and Social Psychology* (106:6), pp. 941-960.

Greitemeyer, T., and Schulz-Hardt, S. 2003a. "Preference-Consistent Evaluation of Information in the Hidden Profile Paradigm: Beyond Group-Level Explanations for the Dominance of Shared Information in Group Decisions," *Journal of Personality and Social Psychology* (84:2), pp. 322-339.

Gunning, R. 1969. "The Fog Index after Twenty Years," *Journal of Business Communication* (6:2), pp. 3-13.

Hall, R. H., and Hanna, P. 2004. "The Impact of Web Page Text-Background Colour Combinations on Readability, Retention, Aesthetics and Behavioural Intention," *Behaviour & information technology* (23:3), pp. 183-195.

Harmon-Jones, E. 2000. "Cognitive Dissonance and Experienced Negative Affect: Evidence That Dissonance Increases Experienced Negative Affect Even in the Absence of Aversive Consequences," *Personality and Social Psychology Bulletin* (26:12), pp. 1490-1501.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., and Merrill, L. 2009. "Feeling Validated Versus Being Correct: A Meta-Analysis of Selective Exposure to Information," *Psychological Bulletin* (135:4), pp. 555-588.

Hastall, M. R. 2009. "Informational Utility as Determinant of Media Choices," in *Media Choice: A Theoretical and Empirical Overview,* T. Hartmann (ed.), New York, NY: Routledge, pp. 149-166.

Hayes, A. F. 2009. "Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium," *Communication Monographs* (76:4), pp. 408-420.

Hayes, A. F., and Rockwood, N. J. 2017. "Regression-Based Statistical Mediation and Moderation Analysis in Clinical Research: Observations, Recommendations, and Implementation," *Behaviour Research and Therapy* (98), pp. 39-57.

He, S. X., and Bond, S. D. 2015. "Why Is the Crowd Divided? Attribution for Dispersion in Online Word of Mouth," *Journal of Consumer Research* (41:6), pp. 1509-1527.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error," *Econometrica: Journal of the Econometric Society* (47:1), pp. 153-161.

Higgins, E. T. 1981. "Role Taking and Social Judgment: Alternative Developmental Perspectives and Processes," in *Social Cognitive Development: Frontiers and Possible Futures*, pp. 119-153.

Hill, B., Ide, S., Ikuta, S., Kawasaki, A., and Ogino, T. 1986. "Universals of Linguistic Politeness: Quantitative Evidence from Japanese and American English," *Journal of Pragmatics* (10:3), pp. 347-371.

Ho, Y.-C., Wu, J., and Tan, Y. 2017. "Disconfirmation Effect on Online Rating Behavior: A Structural Model," *Information Systems Research* (28:3), pp. 626-642.

Hodges, S. D., Clark, B. A. M., and Myers, M. W. 2011. "Better Living through Perspective Taking," in *Positive Psychology as Social Change*, Springer, pp. 193-218.

Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., and Villanueva, B. R. 2010. "Giving Birth to Empathy: The Effects of Similar Experience on Empathic Accuracy, Empathic Concern, and Perceived Empathy," *Personality and Social Psychology Bulletin* (36:3), pp. 398-409.

Hoffman, M. L. 2001. *Empathy and Moral Development: Implications for Caring and Justice*, New York: Cambridge University Press.

Holton, B., and Pyszczynski, T. 1989. "Biased Information Search in the Interpersonal Domain," *Personality and Social Psychology Bulletin* (15:1), pp. 42-51.

Huang, A. H., Chen, K., Yen, D. C., and Tran, T. P. 2015. "A Study of Factors That Contribute to Online Review Helpfulness," *Computers in Human Behavior* (48), pp. 17-27.

Ickes, W., Reidhead, S., and Patterson, M. 1986. "Machiavellianism and Self-Monitoring: As Different as "Me" and "You"," *Social Cognition* (4:1), p. 58.

Ingram, R. E. 1990. "Self-Focused Attention in Clinical Disorders: Review and a Conceptual Model," *Psychological Bulletin* (107:2), p. 156.

Jacowitz, K. E., and Kahneman, D. 1995. "Measures of Anchoring in Estimation Tasks," *Personality and Social Psychology Bulletin* (21:11), pp. 1161-1166.

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

Johnson, B. T. 1994. "Effects of Outcome-Relevant Involvement and Prior Information on Persuasion," *Journal of Experimental Social Psychology* (30:6), pp. 556-579.

Johnson, J. L., Cullen, J. B., Sakano, T., and Takenouchi, H. 1996. "Setting the Stage for Trust and Strategic Integration in Japanese-Us Cooperative Alliances," *Journal of International Business Studies* (27:5), pp. 981-1004.

Jonas, E., Schulz-Hardt, S., Frey, D., and Thelen, N. 2001. "Confirmation Bias in Sequential Information Search after Preliminary Decisions: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information," *Journal of Personality and Social Psychology* (80:4), pp. 557-571.

Jones, Q., Ravid, G., and Rafaeli, S. 2004. "Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* (15:2), pp. 194-210.

Jones, Q., Ravid, G., and Rafaeli, S. 2004. "Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* (15:2), pp. 194-210.

Judd, C. M., Kenny, D. A., and McClelland, G. H. 2001. "Estimating and Testing Mediation and Moderation in within-Subject Designs," *Psychological Methods* (6:2), p. 115.

Keller, P. A., and Block, L. G. 1997. "Vividness Effects: A Resource-Matching Perspective," *Journal of Consumer Research* (24:3), pp. 295-304.

Kim, S.-M., Pantel, P., Chklovski, T., and Pennacchiotti, M. 2006. "Automatically Assessing Review Helpfulness," *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*: Association for Computational Linguistics, pp. 423-430.

Kirmani, A., and Campbell, M. C. 2004. "Goal Seeker and Persuasion Sentry: How Consumer Targets Respond to Interpersonal Marketing Persuasion," *Journal of Consumer Research* (31:3), pp. 573-582.

Kirmani, A., and Campbell, M. C. 2009. "Taking the Target's Perspective: The Persuasion Knowledge Model," in *Social Psychology of Consumer Behavior*, pp. 297-316.

Knobloch, S., Carpentier, F. D., and Zillmann, D. 2003. "Effects of Salience Dimensions of Informational Utility on Selective Exposure to Online News," *Journalism & Mass Communication Quarterly* (80:1), pp. 91-108.

Knobloch-Westerwick, S., and Kleinman, S. B. 2012. "Preelection Selective Exposure," *Communication Research* (39:2), pp. 170-193.

Knobloch-Westerwick, S., Carpentier, F. D., Blumhoff, A., and Nickel, N. 2005. "Selective Exposure Effects for Positive and Negative News: Testing the Robustness of the Informational Utility Model," *Journalism & Mass Communication Quarterly* (82:1), pp. 181-195.

Knobloch-Westerwick, S., Johnson, B. K., and Westerwick, A. 2013. "To Your Health: Self-Regulation of Health Behavior through Selective Exposure to Online Health Messages," *Journal of Communication* (63:5), pp. 807-829.

Koehler, J. J. 1993. "The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality," *Organizational Behavior and Human Decision Processes* (56:1), pp. 28-55.

Kohlberg, L. 1976. "Moral Stages and Moralization: The Cognitive-Developmental Approach," in *Moral Development and Behavior: Theory, Research, and Social Issues,* T. Lickona (ed.), New York: Holt, Rinehart & Winston, pp. 31-53.

Korfiatis, N., GarcíA-Bariocanal, E., and Sánchez-Alonso, S. 2012. "Evaluating Content Quality and Helpfulness of Online Product Reviews: The Interplay of Review Helpfulness Vs. Review Content," *Electronic Commerce Research and Applications* (11:3), pp. 205-217.

Korfiatis, N., Rodriguez, D., and Sicilia, M. A. 2008. "The Impact of Readability on the Usefulness of Online Product Reviews: A Case Study on an Online Bookstore," in *Emerging Technologies and Information Systems for the Knowledge Society,* M.D. Lytras, J.M. Carroll, E. Damiani and R.D. Tennyson (eds.), Berlin: Springer Berlin Heidelberg, pp. 423-432.

Korfiatis, N., Rodríguez, D., and Sicilia, M.-A. 2008. "The Impact of Readability on the Usefulness of Online Product Reviews: A Case Study on an Online Bookstore," in *Emerging Technologies and Information Systems for the Knowledge Society,* M.D. Lytras, J.M. Carroll, E. Damiani and R.D. Tennyson (eds.), Berlin: Springer Berlin Heidelberg, pp. 423-432.

Krishnamoorthy, S. 2015. "Linguistic Features for Review Helpfulness Prediction," *Expert Systems with Applications* (42:7), pp. 3751-3759.

Kronmal, R. A. 1993. "Spurious Correlation and the Fallacy of the Ratio Standard Revisited," *Journal of the Royal Statistical Society. Series A (Statistics in Society)* (156:3), pp. 379-392.

Kumar, N., and Benbasat, I. 2006. "Research Note: The Influence of Recommendations and Consumer Reviews on Evaluations of Websites," *Information Systems Research* (17:4), pp. 425-439.

Kunda, Z. 1990. "The Case for Motivated Reasoning," *Psychological Bulletin* (108:3), pp. 480-498.

Laran, J., Dalton, A. N., and Andrade, E. B. 2011. "The Curious Case of Behavioral Backlash: Why Brands Produce Priming Effects and Slogans Produce Reverse Priming Effects," *Journal of Consumer Research* (37:6), pp. 999-1014.

Lazarsfeld, P., Berelson, B., and Gaudet, H. 1944. *The People's Choice; How the Voter Makes up His Mind in a Presidential Campaign*, New York, NY: Duell Sloan and Pearce.

Lee, M., Rodgers, S., and Kim, M. 2009. "Effects of Valence and Extremity of Ewom on Attitude toward the Brand and Website," *Journal of Current Issues & Research in Advertising* (31:2), pp. 1-11.

Lee, M., Rodgers, S., and Kim, M. 2009. "Effects of Valence and Extremity of Ewom on Attitude toward the Brand and Website," *Journal of Current Issues & Research in Advertising* (31:2), pp. 1-11.

Liang, Y., DeAngelis, B. N., Clare, D. D., Dorros, S. M., and Levine, T. R. 2014. "Message Characteristics in Online Product Reviews and Consumer Ratings of Helpfulness," *Southern Communication Journal* (79:5), pp. 468-483.

Lin, M., Lucas Jr, H. C., and Shmueli, G. 2013. "Research Commentary-Too Big to Fail: Large Samples and the P-Value Problem," *Information Systems Research* (24:4), pp. 906-917.

Liu, B., and Karahanna, E. 2017. "The Dark Side of Reviews: The Swaying Effects of Online Product Reviews on Attribute Preference Construction," *MIS Quarterly* (41:2), pp. 427-448.

Long, E. C., and Andrews, D. W. 1990. "Perspective Taking as a Predictor of Marital Adjustment," *Journal of Personality and Social Psychology* (59:1), pp. 126-131.

Luca, M., and Zervas, G. 2016. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud," *Management Science* (62:12), pp. 3412-3427.

MacKenzie, S. B., and Lutz, R. J. 1989. "An Empirical Examination of the Structural Antecedents of Attitude toward the Ad in an Advertising Pretesting Context," *Journal of Marketing* (53:2), pp. 48-65.

MacKinnon, D. P., Krull, J. L., and Lockwood, C. M. 2000. "Equivalence of the Mediation, Confounding and Suppression Effect," *Prevention Science* (1:4), pp. 173-181.

Mafael, A., Gottschalk, S. A., and Kreis, H. 2016. "Examining Biased Assimilation of Brand-Related Online Reviews," *Journal of Interactive Marketing* (36), pp. 91-106.

Mathieson, A., and Wall, G. 1982. *Tourism, Economic, Physical and Social Impacts*, New York, NY: Longman.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. "An Integrative Model of Organizational Trust," *Academy of Management Review* (20:3), pp. 709-734.

McKinney, V., Yoon, K., and Zahedi, F. M. 2002. "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," *Information systems research* (13:3), pp. 296-315.

Meub, L., and Proeger, T. E. 2015. "Anchoring in Social Context," *Journal of Behavioral and Experimental Economics* (55), pp. 29-39.

Moe, W. W., and Trusov, M. 2011. "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research* (48:3), pp. 444-456.

Montoya, A. K., and Hayes, A. F. 2017. "Two Condition within-Participant Statistical Mediation Analysis: A Path-Analytic Framework," *Psychological Methods* (22:1), pp. 6-27.

Mor, N., and Winquist, J. 2002. "Self-Focused Attention and Negative Affect: A Meta-Analysis," *Psychological Bulletin* (128:4), p. 638.

Mudambi, S. M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com," *MIS Quarterly* (34:1), pp. 185-200.

Mussweiler, T., and Strack, F. 1999. "Hypothesis-Consistent Testing and Semantic Priming in the Anchoring Paradigm: A Selective Accessibility Model," *Journal of Experimental Social Psychology* (35:2), pp. 136-164.

Mussweiler, T., and Strack, F. 2000. "The Use of Category and Exemplar Knowledge in the Solution of Anchoring Tasks," *Journal of Personality and Social Psychology* (78:6), p. 1038.

Mussweiler, T., and Strack, F. 2001. "The Semantics of Anchoring," *Organizational Behavior and Human Decision Processes* (86:2), pp. 234-255.

Mussweiler, T., Strack, F., and Pfeiffer, T. 2000. "Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility," *Personality and Social Psychology Bulletin* (26:9), pp. 1142-1150.

Newcomb, T. M. 1956. "The Prediction of Interpersonal Attraction," *American Psychologist* (11:11), p. 575.

Nickerson, R. S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of general psychology* (2:2), p. 175.

Nisbett, R. E., and Ross, L. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs, NJ: Prentice-Hall.

Papke, L. E., and Wooldridge, J. 1993. "Econometric Methods for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates." National Bureau of Economic Research Cambridge, Mass., USA.

Pease, M. E., Brannon, L. A., and Pilling, V. K. 2006. "Increasing Selective Exposure to Health Messages by Targeting Person Versus Behavior Schemas," *Health Communication* (19:3), pp. 231-240.

Pennebaker, J. W. 2011. *The Secret Life of Pronouns: What Our Words Say About Us*, New York: Bloomsbury Press.

Pennebaker, J. W., and Lay, T. C. 2002. "Language Use and Personality During Crises: Analyses of Mayor Rudolph Giuliani's Press Conferences," *Journal of Research in Personality* (36:3), pp. 271-282.

Pennebaker, J. W., Booth, R. J., and Francis, M. E. 2007. "Liwc2007: Linguistic Inquiry and Word Count," in: *Austin, Texas: liwc. net*.

Pennebaker, J. W., Mehl, M. R., and Niederhoffer, K. G. 2003. "Psychological Aspects of Natural Language Use: Our Words, Our Selves," *Annual Review of Psychology* (54:1), pp. 547-577.

Petrocelli, J. V., Tormala, Z. L., and Rucker, D. D. 2007a. "Unpacking Attitude Certainty: Attitude Clarity and Attitude Correctness," *Journal of personality and social psychology* (92:1), p. 30.

Petrocelli, J. V., Tormala, Z. L., and Rucker, D. D. 2007b. "Unpacking Attitude Certainty: Attitude Clarity and Attitude Correctness," *Journal of Personality and Social Psychology* (92:1), pp. 30-41.

Petrova, P. K., and Cialdini, R. B. 2005. "Fluency of Consumption Imagery and the Backfire Effects of Imagery Appeals," *Journal of Consumer Research* (32:3), pp. 442-452.

Petty, R. E., Briñol, P., Tormala, Z. L., and Wegener, D. T. 2007. "The Role of Meta-Cognition in Social Judgment," in *Social Psychology: Handbook of Basic Principles,* H. ET and K. AW (eds.), New York: Guilford Press, pp. 254-284.

Phillips, O. R., and Menkhaus, D. J. 2010. "The Culture of Private Negotiation: Endogenous Price Anchors in Simple Bilateral Bargaining Experiments," *Journal of Economic Behavior & Organization* (76:3), pp. 705-715.

Piaget, J. 1932. *The Moral Development of the Child*, London: Kegan Paul, Trench & Trübner.

Plank, R. E., Minton, A. P., and Reid, D. A. 1996. "A Short Measure of Perceived Empathy," *Psychological Reports* (79:3 suppl), pp. 1219-1226.

Pornpitakpan, C. 2004. "The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence," *Journal of Applied Social Psychology* (34:2), pp. 243-281.

Price, L. L., Feick, L. F., and Guskey, A. 1995. "Everyday Market Helping Behavior," *Journal of Public Policy & Marketing* (14:2), pp. 255-266.

Prislin, R., and Wood, W. 2005. "Social Influence in Attitudes and Attitude Change," in *The Handbook of Attitudes,* D. Albarracín, B.T. Johnson and M.P. Zanna (eds.), Hillsdale, NJ: Erlbaum, pp. 671-706.

Qiu, L., Pang, J., and Lim, K. H. 2012. "Effects of Conflicting Aggregated Rating on Ewom Review Credibility and Diagnosticity: The Moderating Role of Review Valence," *Decision Support Systems* (54:1), pp. 631-643.

Rholes, W. S., Simpson, J. A., Tran, S., Martin, A. M., and Friedman, M. 2007. "Attachment and Information Seeking in Romantic Relationships," *Personality and Social Psychology Bulletin* (33:3), pp. 422-438.

Roberts, J. H., and Lattin, J. M. 1991. "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research* (28:4), pp. 429-440.

Rozin, P., and Royzman, E. B. 2001. "Negativity Bias, Negativity Dominance, and Contagion," *Personality and social psychology review* (5:4), pp. 296-320.

Rude, S., Gortner, E.-M., and Pennebaker, J. 2004. "Language Use of Depressed and Depression-Vulnerable College Students," *Cognition & Emotion* (18:8), pp. 1121-1133.

Russo, J. E., Meloy, M. G., and Medvec, V. H. 1998. "Predecisional Distortion of Product Information," *Journal of Marketing Research* (35:4), pp. 438-452.

Sagarin, B. J., Cialdini, R. B., Rice, W. E., and Serna, S. B. 2002. "Dispelling the Illusion of Invulnerability: The Motivations and Mechanisms of Resistance to Persuasion," *Journal of Personality and Social Psychology* (83:3), p. 526.

Sargent, S. L. 2007. "Image Effects on Selective Exposure to Computer-Mediated News Stories," *Computers in Human Behavior* (23:1), pp. 705-726.

Sassenrath, C., Hodges, S. D., and Pfattheicher, S. 2016. "It's All About the Self: When Perspective Taking Backfires," *Current Directions in Psychological Science* (25:6), pp. 405-410.

Schindler, R. M., and Bickart, B. 2012. "Perceived Helpfulness of Online Consumer Reviews: The Role of Message Content and Style," *Journal of Consumer Behaviour* (11:3), pp. 234-243.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., and Simons, A. 1991. "Ease of Retrieval as Information: Another Look at the Availability Heuristic," *Journal of Personality and Social psychology* (61:2), p. 195.

Sears, D. O. 1965. "Biased Indoctrination and Selectivity of Exposure to New Information," *Sociometry* (28:4), pp. 363-376.

196

Sen, S., and Lerman, D. 2007. "Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing* (21:4), pp. 76-94.

Sen, S., and Lerman, D. 2007. "Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing* (21:4), pp. 76-94.

Shocker, A. D., Ben-Akiva, M., Boccara, B., and Nedungadi, P. 1991. "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions," *Marketing Letters* (2:3), pp. 181-197.

Shoham, M., Moldovan, S., and Steinhart, Y. 2017. "Positively Useless: Irrelevant Negative Information Enhances Positive Impressions," *Journal of Consumer Psychology* (27:2), pp. 147-159.

Simmons, R. A., Gordon, P. C., and Chambless, D. L. 2005. "Pronouns in Marital Interaction What Do "You" and "I" Say About Marital Health?," *Psychological Science* (16:12), pp. 932-936.

Simon, H. A. 1990. "Invariants of Human Behavior," *Annual review of psychology* (41:1), pp. 1-20.

Smith, R. E., and Swinyard, W. R. 1988. "Cognitive Response to Advertising and Trial: Belief Strength, Belief Confidence and Product Curiosity," *Journal of Advertising* (17:3), pp. 3-14.

Snyder, M. 1981. "On the Self-Perpetuating Nature of Social Stereotypes," in *Cognitive Processes in Stereotyping and Intergroup Behavior,* D.L. Hamilton (ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 183 - 212.

Snyder, M. 1984. "When Belief Creates Reality," *Advances in Experimental Social Psychology* (18), pp. 247-305.

Sparks, B. A., Perkins, H. E., and Buckley, R. 2013. "Online Travel Reviews as Persuasive Communication: The Effects of Content Type, Source, and Certification Logos on Consumer Behavior," *Tourism Management* (39), pp. 1-9.

Strack, F., and Mussweiler, T. 1997. "Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility," *Journal of Personality and Social Psychology* (73:3), p. 437.

Stroud, N. J. 2008. "Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure," *Political Behavior* (30:3), pp. 341-366.

Sun, M. 2012. "How Does the Variance of Product Ratings Matter?," *Management Science* (58:4), pp. 696-707.

Taber, C. S., and Lodge, M. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science* (50:3), pp. 755-769.

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., and Galinsky, A. D. 2011. "Perspective Taking Combats Automatic Expressions of Racial Bias," *Journal of Personality and Social Psychology* (100:6), p. 1027.

Toi, M., and Batson, C. D. 1982. "More Evidence That Empathy Is a Source of Altruistic Motivation," *Journal of Personality and Social Psychology* (43:2), pp. 281-292.

Tversky, A., and Kahneman, D. 1973. "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive psychology* (5:2), pp. 207-232.

Tversky, A., and Kahneman, D. 1974. "Judgment under Uncertainty: Heuristics and Biases," *science* (185:4157), pp. 1124-1131.

Vorauer, J. 2013. "The Case for and against Perspective-Taking," *Advances in Experimental Social Psychology* (48), pp. 59-115.

Wansink, B., Kent, R. J., and Hoch, S. J. 1998. "An Anchoring and Adjustment Model of Purchase Quantity Decisions," *Journal of Marketing Research*), pp. 71-81.

Wegener, D. T., Petty, R. E., Blankenship, K. L., and Detweiler- Bedell, B. 2010. "Elaboration and Numerical Anchoring: Implications of Attitude Theories for Consumer Judgment and Decision Making," *Journal of Consumer Psychology* (20:1), pp. 5-16.

Weintraub, W. 1989. *Verbal Behavior in Everyday Life*, New York:: Springer.

Williams, P., Fitzsimons, G. J., and Block, L. G. 2004. "When Consumers Do Not Recognize "Benign" Intention Questions as Persuasion Attempts," *Journal of Consumer Research* (31:3), pp. 540-550.

Wispé, L. 1986. "The Distinction between Sympathy and Empathy: To Call Forth a Concept, a Word Is Needed," *Journal of Personality and Social Psychology* (50:2), p. 314.

Woodside, A. G., and MacDonald, R. 1994. "General System Framework of Customer Choice Processes of Tourism Services," in *Spoilt for Choice,* R.V. Gasser and K. Weiermair (eds.), Thaur, Germany: Kulturverlag.

Wooldridge, J. M. 2011. "Fractional Response Models with Endogeneous Explanatory Variables and Heterogeneity," *CHI11 Stata Conference*: Stata Users Group.

Wright, P., and Barbour, F. 1977. *Phased Decision Strategies: Sequels to an Initial Screening*, Graduate School of Business, Stanford University.

Wyer, R. S., and Albarracín, D. 2005. "The Origins and Structure of Beliefs and Goals," in *Handbook of Attitudes and Attitude Change,* D. Albarracín, B.T. Johnson and M.P. Zanna (eds.), Hillsdale, NJ: Erlbaum, pp. 273-322.

Wyer, R. S., and Frey, D. 1983. "The Effects of Feedback About Self and Others on the Recall and Judgments of Feedback-Relevant Information," *Journal of Experimental Social Psychology* (19:6), pp. 540-559.

Yang, Z., Cai, S., Zhou, Z., and Zhou, N. 2005. "Development and Validation of an Instrument to Measure User Perceived Service Quality of Information Presenting Web Portals," *Information & Management* (42:4), pp. 575-589.

Yin, D., Bond, S. D., and Zhang, H. 2014. "Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews," *MIS Quarterly* (38:2), pp. 539-560.

Yin, D., Bond, S. D., and 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), pp. 447-463.

Yin, D., Mitra, S., and Zhang, H. 2016. "Research Note—When Do Consumers Value Positive Vs. Negative Reviews? An Empirical Investigation of Confirmation Bias in Online Word of Mouth," *Information Systems Research* (27:1), pp. 131-144.