DigitalCommons @WayneState

Wayne State University

Wayne State University Dissertations

1-1-2017

Star-Crossed Consumers: The Effects Of Online Rating Scale Length On Product Evaluations

Aaron C. Johnson
Wayne State University,

Follow this and additional works at: https://digitalcommons.wayne.edu/oa_dissertations

Part of the Marketing Commons

Recommended Citation

Johnson, Aaron C., "Star-Crossed Consumers: The Effects Of Online Rating Scale Length On Product Evaluations" (2017). *Wayne State University Dissertations*. 1816.
https://digitalcommons.wayne.edu/oa_dissertations/1816

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.

STAR-CROSSED CONSUMERS: THE EFFECTS OF ONLINE RATING SCALE LENGTH ON PRODUCT EVALUATIONS

by

AARON JOHNSON

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2017

MAJOR: BUSINESS ADMIN (Marketing)

Approved By:

____

Advisor Date

____

Advisor Date

____

____

____

© COPYRIGHT BY

AARON JOHNSON

2017

All Rights Reserved

DEDICATION

To my wife, Jen, and daughters, Jane and Nora: Without your love and support, this journey would have been far less rewarding.

ACKNOWLEDGEMENTS

To my advisors, Dr. Abhijit Biswas and Dr. Sujay Dutta: This accomplishment would have been impossible without your continued teaching and leadership. Thank you for the countless hours you spent to instill in me the importance of critical thinking and academic rigor. Thank you also for your friendship and humor that made this such a memorable experience.

To my unofficial advisors, Dr. Attila Yaprak, Dr. Andrea Tangari and Dr. David Merolla: Thank you for always keeping your offices open to me and lending an ear to my many inquiries. Your kindness and guidance have served me well and will not be forgotten.

A special thanks to my marketing friends, Ahmet, Nick, Somak and Swati. Sharing this experience with you has been wonderful and truly makes this milestone in my life bittersweet.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The purpose of this research is to investigate the effects of online product rating scale length on consumer perceptions of product quality and purchase intention. Consumers can review ratings of individuals who have evaluated a product, and because of this, online product ratings offer valuable insights in a pre-purchase setting. Specifically, we analyze differences in online ratings by comparing 5- and 10-point rating scale lengths, which are commonly seen in the online market. We begin the introduction by discussing asymmetric information as a preliminary motive for consumers to openly share information about product quality. Information sharing can reduce search costs as consumers can rely on the opinion of others to better determine a quality product that is right for them. Traditionally, individuals overcame such concerns through face-to-face word-of-mouth communication methods. Electronic word-of-mouth naturally followed as the internet allowed for universal communication. Next, we discuss how electronic word-of-mouth has been adopted by companies and consumers alike to show its growing influence in the market. Lastly, we outline the scope of the present research in more detail and provide a review of subsequent chapters in this research.

Word-of-Mouth

Asymmetric Information and Search Costs

The central problem of consumer behavior is choice (Taylor 1974). This issue may result from both the inherent costs of searching for a product that is appropriately priced and the quality information asymmetry between a buyer and a seller (Boulding and Kirmani, 1993). Communication among individuals is a time-honored method of alleviating purchasing uncertainties, as humans commonly share their consumption experiences with close friends or family. In recent years, online product review networks have provided additional help by diffusing vast amounts of information that ecommerce customers can use to more easily evaluate products. Such modern methods of information sharing may reduce the dissonance from information asymmetry and the costs of searching for product information to evaluate the price and quality of different products.

Chen et al. (2006) claim that consumers encounter search costs in seeking for product quality information and identifying a product that "fits" with their consumption tastes. Obviously, ecommerce can help to alleviate costs associated with imperfect information about prices, because click-of-a-mouse shopping has seemingly enhanced the consumer experience. However, the costs associated with quality and fit may be less clearly offset by ecommerce, but given online consumer rating and review platforms there are now mountains of online information to better determine appropriate product quality, not to mention price and product fit, to guide purchasing behavior like never before.

The proverbial "kicking the tires" may still exist for consumers in the automotive market, yet a sea of other products and services are now purchased from the comfort of a home computer or handheld device without ever handling the product prior to purchase. The day and age of armchair consumers is changing the way buyers and sellers communicate and exchange in the market. Due to a lack of physical quality cues in online markets, most retailers provide online rating and review systems within their websites for consumers to share their experiences about the quality of purchased products. Again, the process of openly sharing information provides the potential for consumers to reduce uncertainty about product quality (Dellarocas, 2003).

Traditional Word of Mouth (WOM)

Social pressure influences our decision-making and purchase behavior. Despite how it may be portrayed to children, this social phenomenon encompasses more than negative influence. Since the beginning of human society, WOM has been a powerfully influential source of information transmission (Godes and Mayzlin 2004). The authors further argue that there is good reason to believe that WOM has more potential impact than any other communication channel.

In the past, individual consumers accessed only a small sample of others' evaluations, typically through traditional communication channels like friends, family members, and critics (He and Bond 2015). In a consumption context, consumers have relied on WOM to mitigate search costs and asymmetric information. There is empirical support, as well as an intuitive reasoning, for the proposed link between WOM and consumer behavior (Godes and Mayzlin 2004). WOM has been studied for decades, but has arguably become more relevant with the arrival of the internet.

Electronic Word of Mouth (eWOM)

Henning-Thurau et al (2004) describe eWOM as any statement made by customers about a product or company, which is made available via the Internet. This new form of consumer communication has received increased attention from researchers (Zhang, Craciun et al. 2010). Research findings support the notion that online consumer product ratings can act as a reliable summary for overall sentiment that may be expressed in written consumer reviews. (Zhu & Zhang, 2010), which, as previously mentioned, can have a strong influence on the decision-making processes of other online shoppers. An electronic forum provides an environment where individuals are often exposed to an incredible range of opinions, across a broad diversity of sources. Compared to traditional WOM, where consumers seek and share opinions within a small circle of influence, online communications have provided consumers access to the opinions of thousands of strangers (He and Bond 2015). Thus, it seems that when an opinion is expressed by more people it is more difficult to ignore (Khare, Labrecque et al. 2011). The advances of information technology have profoundly changed the way information is transmitted and have transcended the traditional limitations of WOM (Laroche et al. 2005). Traditional WOM has been transformed into permanent online messages visible to a world audience, and, as a result, eWOM plays a significant role in consumer purchase decisions (Duan, Gu et al. 2008).

As early as 1991, Bakos argued that with the introduction of electronic markets, the power structure between buyers and sellers will shift in favor of buyers. This idea was later supported by Rust and Oliver (1994) who predicted that online communication would dramatically increase the quantity and quality of information available to the consumer. It is possible that we are observing a historic transition of power-one that transfers power from the mightiest corporations and gives it to consumers (Murphy, 2000). It would seem that the internet has singlehandedly perpetuated these contemporary procedures of shopping and information sharing, which arguably has shifted power from the firm to consumers.

Online consumer reviews (OCRs) have become an integral part of the decision-making process for online consumers. OCRs can comprise quantitative ratings and qualitative reviews regarding product evaluations. Online ratings and reviews allow individuals to make post-purchase evaluation of a product and summarize their overall experience by means of an average numerical scale rating (summary rating) and/or a text-based review, for example. Consumers can use OCRs to virtually find the best price-quality combination (Shipman, 2001). Valence is often displayed as non-numeric symbols (e.g., stars) that are commonly used to by critics or professional raters to assess certain retail establishments, such as hotels and restaurants. With the advent of evaluation interfaces (e.g., Apple's App Store or Amazon.com.), star ratings have become a ubiquitous way to rate products of every category imaginable. Table 1.1 provides a list of retailers and review sites that provide online product ratings. The table is organized by the site or retailer that is hosting the rating interface, if stars are used to showcase the rating scale, and if valence (summary rating), volume (number of users who provided ratings), or variance (rating scale distribution) are displayed. Lastly, we record the scale length used to capture consumer evaluations (e.g., 5-point or 10-point). Although this list is not exhaustive, it offers a glimpse into some of the methods for displaying product evaluative ratings. Per Table 1, 59% of the sites use star symbols to visually represent consumer ratings and 60% display the summary rating alongside the stars. Furthermore, 72% of the rating platforms in the list use a 5-point rating scale, while only 9% use a 10-point scale length.

Table 1: OCR Rating Platforms in the eMarket

Site/Retailer	Stars	Valence	Volume	Variance	Scale
General Retailers (online only)					
Amazon	y	y	y	y	5
eBay	y	y	y	y	5
Zappos	y	n	y	y	5
Etsy	y	n	y	n	5
Overstock	y	y	y	y	5
Google	y	y	y	n	5
General Retailer (in-store)					
Walmart	y	y	y	y	5
Costco	y	y	y	y	5
Sam's Club	y	n	y	y	5
Sears	y	y	y	y	5
Target	y	y	y	y	5
Home					
Home Depot	y	y	y	y	5
Lowes	y	n	y	y	5
Menards	n	n	n	n	na
Ace Hardware	y	y	y	n	5
Rent	y	n	y	y	5
Renters Voice	y	n	y	n	5
Electronics					
Best Buy	y	y	y	y	5
Radio Shack	y	n	n	n	5

Dell	y	y	y	n	5
Staples	y	y	y	y	5
Apple	n	n	n	n	na
Engadget	n	y	n	n	100
Department Stores					
Old Navy	y	y	y	y	5
Kohl's	y	y	y	n	5
Gap	y	y	y	y	5
Banana Republic	y	y	y	y	5
Macys	y	n	y	y	5
Nordstrom	y	y	y	y	5
Neiman Marcus	n	n	n	n	na
Saks Fifth Ave	y	y	y	n	5
Bloomingdales	n	y	y	n	5
Lord and Taylor	y	y	y	n	5
Barneys New York	n	n	n	n	na
Gucci	n	n	n	n	na
Burberry	n	n	n	n	na
Tiffany and Co.	n	n	n	n	na
Dolce and Gabbana	n	n	n	n	na
Food & Entertainment					
Apple App Store	y	n	y	y	5
IMDb	n	y	y	y	10
Good Reads	y	y	y	y	5
Yelp	y	n	y	y	5
Zomato (urbanspoon)	n	y	y	n	5
Rotten Tomatoes	n	y	y	y	5, 10
Zagat	n	y	n	n	5
GameStop	n	y	y	y	10
Metacritic	n	y	y	y	100
Flixster	n	n	n	n	100
Netflix	n	y	y	y	5, 10
Groupon	y	y	y	n	5
PCmag	n	na	na	na	5
Travel					
Orbitz	n	y	y	y	5
Expedia	n	y	y	n	5
Trip Advisor	n	n	y	n	5
Priceline	n	y	y	y	10
Hipmunk	n	y	y	n	10
Hotwire	n	y	y	n	10
Travelocity	n	y	y	n	5
Booking	n	y	y	y	10
Hotels	n	y	y	y	5

Uber	y	y	y	y	n
AirBnB	y	n	y	n	5
Edmunds	y	n	y	y	5
JD Power	n	y	na	y	5
Cars	y	y	y	n	5
Dealer Rater	y	y	y	n	5
KBB					
Personal					
Health Grades	y	y	y	n	5
Rate MDs	y	n	y	n	5
Vitals	y	y	y	y	5
Career Bliss	y	n	y	n	5
Rate My Employer	y	n	y	n	5
Job Advisor	n	y	y	n	5
Rate My Professor	n	y	y	y	5
AVVO	y	n	y	n	5
Mechanic Ratingz	n	y	y	n	5
Miscellaneous					
Angie's List	n	n	y	n	Letter
Consumer Reports	n	n	n	n	na
City Search	n	n	y	n	%
CNET	y	n	y	y	5
BBB	n	na	na	na	Letter
Insider Pages	y	n	y	n	5
Judy Book	y	y	y	y	5
Merchant Circle	y	y	y	n	5
Yellow Pages	y	n	y	n	5
Indeed	y	y	y	y	5
Kununu	y	y	y	y	5
ePinions	y	n	y	n	5

*Stars (star symbol rating present); Valence (summary rating displayed); Volume (number of ratings displayed); Variance (ability to view rating distribution); Scale (length of rating scale)

It is widely accepted that there has been a dramatic surge in the volume and general availability of online reviews, now often called "word of mouse" (Clemons, Gao et al. 2006). The Web has become a tremendously efficient medium to grasp a universal market, regardless of geographic boundaries (Duan, Gu et al. 2008). Due to the openness and connectivity of the Internet, OCRs are being generated at an unprecedented scale and speed (Wu 2013).

Researchers have studied the effects of OCRs on buyer behavior using many factors. Some of the effects of OCRs have been noted by authors such as Chen (2011), who posits that brand names will lose much of their importance in the interactive marketing environment, suggesting that consumers will bypass marketer-influenced quality signals and instead rely more on user-generated OCRs. Because OCRs do not originate with the company, it is considered highly credible and influential (Bickart and Schindler 2001). Additionally, Chen and Xie (2008) highlight the effect of OCRs on novice consumers' identification of products that best match their preferences. In the absence of text review or summary rating information, novice consumers may be less likely to buy a product if only seller-created product attribute information is available, suggesting that the availability of OCRs may lead to an increase in sales (Moe and Trusov 2011; Chen, 2008). In any case, these few examples of eWOM factors, and many others, play an important role in determining how, when and why online consumption happens and is of much interest to the firm.

Word-of-Mouth in the eMarket

eWOM Usage by the Firm

OCR represents a potentially valuable tool for firms, who can use them to monitor consumer attitudes toward their products and adapt their marketing practices accordingly (Dellarocas, Zhang et al. 2007). Firms are interested in eWOM communication because it affects consumers' willingness to pay for products and product sales (Chen and Lurie 2013). OCRs are available for everything from books (nybooks.com), cameras (www.dpreview.com), and movies (mrqe.com), to consumer electronics (cnet.com), travel (tripadvisor.com), and beer (beerhunter.com) (Clemons, Gao et al. 2006). As an unpaid endorsement for products or services, WOM is perhaps the most believable form of advertising for marketers (Henricks, 1998).

Despite the idea that user-generated content is unfiltered, compared to company-generated product information, firms recognize the critical role it plays in sales. Online sites like Amazon.com, with their endless supply of products and OCR, have optimized the shopping experience and seemingly minimized consumer’s search costs. Even traditional in-store retail establishments like Wal-Mart and Target now include a vast online inventory to meet the demands of virtual shoppers. Amazon has even eliminated its television and general-purpose print advertising budgets as a result of OCRs (Sen and Lerman 2007). The firm believes that its consumers trust other consumers’ opinions more than they do traditional advertising, and that such eWOM is more effective in influencing consumer behavior (Thompson 2003).

Many e-commerce companies, such as Amazon and eBay, both solicit and publish customers’ opinions about the products they have purchased. These, and many other firms, are taking advantage of OCRs as a new marketing tool (Dellarocas 2003). Studies show that firms not only regularly sponsor promotional chats on online forums, such as USENET (Mayzlin 2006), but also proactively induce their consumers to spread the word about their products online (Zhu and Zhang 2010).

Given the immense number of opinions available, it is common for online platforms to summarize evaluations in graphical form, such as making the variance (rating distribution) available. As a result, these summaries may play an increasingly important role in consumer decision-making (He and Bond 2015). The literature has identified an overwhelming amount of OCR that exist in eWOM, so it stands to reason why firms summarize these quantitative ratings for their consumers.

eWOM Usage by the Consumer

Given the ubiquity of star-ratings and text-based reviews in ecommerce (e.g., Amazon) and crowd-source review sites (e.g., Yelp) OCRs play an increasingly significant role in consumer purchase decisions. A 2007 survey by comScore found that 75% of consumers are making use of product ratings and text reviews before purchasing products online, and 24% of internet users even access OCRs prior to paying for a service delivered offline (Zhu and Zhang 2010). Researchers have identified the use of OCR as a common step for today's consumer, and supported these claims with statistics. Although these reported statistics are often mixed, OCRs are decidedly used in pre-purchase evaluations of a product. Below we will discuss some of the reported figures.

OCR communication is highly trusted by online shoppers and over 60% of consumers consult online reviews before making buying decisions (Chen and Lurie 2013). Fagerstrom, Ghinea et al. (2016) reported on a survey by Forrester Research based on more than 58,000 U.S. respondents which found that approximately 70% of online customers rely on brand or product recommendations from friends and family, whereas 46% of the respondents reported that they rely on consumer-written online reviews. Citing from the same survey, another article reported that 64% of the respondents want to see user ratings and reviews on the e-commerce websites they visit (Sun 2012). Additionally, Schlosser (2011) suggests that 58% of consumers prefer sites with peer reviews and nearly all (98%) online shoppers reported reading peer reviews before making a purchase. Anderson (2014) revealed that 88% of consumers trust OCR as much as personal recommendations and 85% of them read up to 10 reviews whenever they want to shop online.

The reported levels of OCR usage seem to vary but, overall, seem substantial. OCRs, via numerical ratings and text reviews, are prevalent in online consumption. Thus, the conclusion that OCR has become a critical source of information for consumers regarding product quality is intuitively understandable (Decker and Trusov 2010). Apparently the seemingly majority of online consumers value the opinions of others, but to what extant? It remains unclear how a quantitative summary rating is viewed in comparison to text-based reviews while consumers evaluate a product.

Text-based Review Complexity

Imagine an online purchase scenario, where a consumer clicks to view a product of interest. Typically, next to the product image is the quantitative summary rating (usually depicted by colored stars or a numerical value). In addition, the quantitative summary rating might be accompanied by the number of raters and sometimes the distribution of ratings. Usually, by clicking on one of these summary ratings or by scrolling down further on the webpage, consumers can only then read text reviews. Quantitative summary ratings seem to be the face of the OCR experience, whereas text-based reviews come last in the typical evaluation sequence. This alone does not discount the qualitative effect of written opinions, but should give pause to researchers who discount quantitative summary ratings because of their simplicity. Despite the abundant presence of data in the form of text-based reviews, we seek to stress the importance of summary ratings.

The OCR literature of the past decade focuses primarily on text-based reviews (e.g., Basuroy et al. 2003; Huang and Chen, 2006; Ludwig et al. 2013). As noted by Resnick et al. (2000), numerical ratings fail to convey the important subtleties of online interactions, like the reputations of the people providing the feedback. Chen (2006), for example, examines the role of reviewer reputation, where reputation is based on how other online users evaluate an individual's written review. De Maeyer (2012) adds that most rating distributions are bimodal, and that the only way consumers can make sense of this is to read the text reviews. Summary rating findings are sometimes inconclusive or conflicting (Dellarocas, 2003; Hu, Liu, & Zhang, 2008), so it is unclear what factors truly influence summary ratings. Yet, it may not be appropriate to brush over these simple summary measures and only read text reviews, as the written word can be convoluted. A focus on text-based review research assumes that consumers deep dive for more detailed information. Furthermore, it assumes that they are motivated and able to perform this cognitively elaborate task of evaluating qualitative opinions.

The difficulty in processing text-based reviews offers support to the importance of quantitative summary ratings. Cognitive elaboration literature states that people may generally lack the ability, and motivation to think critically about information (Petty et al. 1997). Even for experienced online shoppers, they have limited time and cognitive resources to devote to the dozens, hundreds, or thousands of written reviews before deciding. Although text reviews allow for a more complete analysis of positive and negative sentiment, it requires extra time and processing abilities that consumers may lack. Text-based reviews require more ability and motivation to read and process than summary ratings, especially with a high volume of text reviews. Thus, consumers may be less likely to engage in more difficult elaborative task of filtering through, and reading, text reviews.

Consumers have limited time and cognitive resources to read enough negative and positive text reviews to gain a balanced understanding of the expressed consumer sentiment. Although Mousavizadeh, Koohikamali et al. (2015) state that people read up to 10 reviews before purchasing a product online, this research fails to identify the proportion of negative to positive ratings, or how these text-based reviews interact with summary ratings. Furthermore, spotlight reviews (those listed first) have been shown to have a larger positive marginal impact on sales than other reviews (Chen et al. 2006), analogous to a top-page search result bias using an online search engine. Although text sentiment allows consumers to better explain their experience with the product, both the pros and cons, it is not organized as efficiently as summary ratings. Although summary ratings contain less sentiment than text reviews, it can be argued that they simply and effectively posit an average of the general sentiments of all respondents.

Scope of Present Research

The overall context of this study pertains to quantitative summary ratings and not text reviews. Researchers have devoted much attention to the rich content in text-based sentiment, yet we argue that summary ratings are also a critical aspect of e-commerce that is worthy of research. Although there is a rich stream of research regarding summary ratings there appears to be a void in explaining how consumption behavior differs when viewing summary ratings originating from different rating scale lengths. The object of our research is not to investigate how consumers rate products using different scale lengths, but how consumers perceive the ratings of the already-rated products.

Our search of the online market demonstrates that most sites containing summary ratings are 5-star in length and those that employ a 10-star rating scale are typically used for experience goods, like movies and hotels. WOM has been frequently cited as the single most important factor that determines the long-term success of experience goods (Godes and Mayzlin 2004), but we also want to see what the effect will be between 5- and 10-point scale lengths for a search good.

We show across three studies how the effect of scale length on perceived product quality and purchase intention is statistically stronger for the 10-point summary rating scale length, compared to a 5-point scale length of equal proportion. Potential moderating variables are included in Study Two (presence of rating percentage) and Study Three (various levels of consumer rating volume) to explore boundary conditions of the effect of rating scale length. We will use heuristic processing via an anchoring mechanism to explain the effect on purchase intention and perceived product quality between 5- and 10-point summary rating scale lengths.

The rest of this dissertation progresses as follows. Chapter 2 provides a review of the relevant theories used and the findings from existing literature related to quantitative summary ratings. Chapter 3 presents the current research questions in more detail. Chapter 4 is a discussion of the conceptual background and development of our key hypotheses on the differential effects of scale lengths. Chapter 5 presents an experiment designed and conducted to test the key hypotheses. Chapters Six and Seven present the design, procedures, and results of two experimental studies designed and conducted to test probable boundary conditions. Finally, Chapter 8 presents a discussion of the findings, and limitations of the present research, together with directions for future research.

CHAPTER 2: LITERATURE REVIEW

OCRs can be classified as either qualitative or quantitative (Sridhar & Srinivasan, 2012). Qualitative OCRs provide a text-based review, or sometimes even a video, of the consumer's usage experience. In such qualitative reviews, individuals can describe, criticize, and evaluate the product (Kostyra et al. 2016). In the case of a quantitative OCR, the customer typically provides a single rating to summarize their product evaluation. Quantitative OCRs from individual consumers are then aggregated into a summary statistic to be displayed on a webpage as a valence. Summary statistics allow consumers to assess product quality more easily instead of filtering through each individual rating or review. Since the focus of our paper is to test the effects of quantitative summary measures, we will first define the important quantitative OCR factors that are discussed and tested in the literature. Per Chintagunta et al. (2010), a quantitative OCR comprises the following three factors:

1. Valence: an average numerical customer satisfaction rating, which we mainly refer to as a summary rating in other chapters of this research. For example, an online rating platform may identify a numerical product rating, such as 4 out of 5, or 8 out of 10. Often, this numerical rating is displayed as a single value (e.g., either "4" or "8") next to colored stars to clarify the rating percentage average.

2. Volume: the total number of customer ratings. For example, an online rating of 4 out of 5 will also include another numerical value, signifying the number of customer ratings that contributed to the summary rating.

3. Variance: the variation in customer ratings along the rating scale. For example, after viewing an online rating of 3.75 out of 5, with 100 consumer ratings, consumers may additionally view a distribution of the ratings for each of the 5 scale points: 0 ratings at 1 out of 5; 15 ratings at 2 out

of 5; 0 ratings at a 3 out of 5; 80 ratings at 4 out of 5; and 5 ratings at a 5 out of 5. The total score of rating points is 375, and when divided by the total number of ratings (100) the summary rating equals 3.75.

Company sites differ in their variations of displaying summary ratings, volume, and variance, as seen in Table 1.1. Although we are focusing on quantitative OCR instead of text-based reviews, it is difficult to completely isolate these quantitative factors, as these rating and text measures are intimately connected. Much of the literature that focuses primarily on quantitative ratings often explores the interactive effects of elements relating to qualitative OCR. Our study will not include any text-based review variables, yet the literature that we summarize occasionally includes both quantitative and qualitative OCR variables, including several other explanatory moderators and unique dependent variables (see Figure 2.1). Most of the previous literature has investigated summary ratings and volume, with only a few studies considering the effects of the variance or the interaction of these OCR variables. Furthermore, valence and volume effects are predominately found to be positive, with far less support for negative or non-significant effects. Compiling these mixed results will provide a better overall understanding of the various relationships between these OCR variables and marketing measures of interest. Often, mixed results confirm the necessity to study summary ratings in more detail and explore moderating factors to explain the nuances in quantitative OCRs.

The remainder of this chapter will review the OCR literature, primarily in the context of quantitative measures, by examining findings in terms of valence, volume, variance and their interactions. Interactions between the OCR quantitative terms themselves (e.g., valence and volume) and interactions between the OCR quantitative terms and another unique variable (e.g., valence and price) will also be discussed within each appropriate section and not separately.

Furthermore, our review of the existing literature will examine the valence, volume, and variance variables individually, and not by author(s). Both marketing and non-marketing literature was utilized in this review, spanning multiple disciplines. To view the overall findings of each quantitative OCR variable by article, please see Table 2.1.

Figure 2: Combined Models of Quantitative OCR Studies

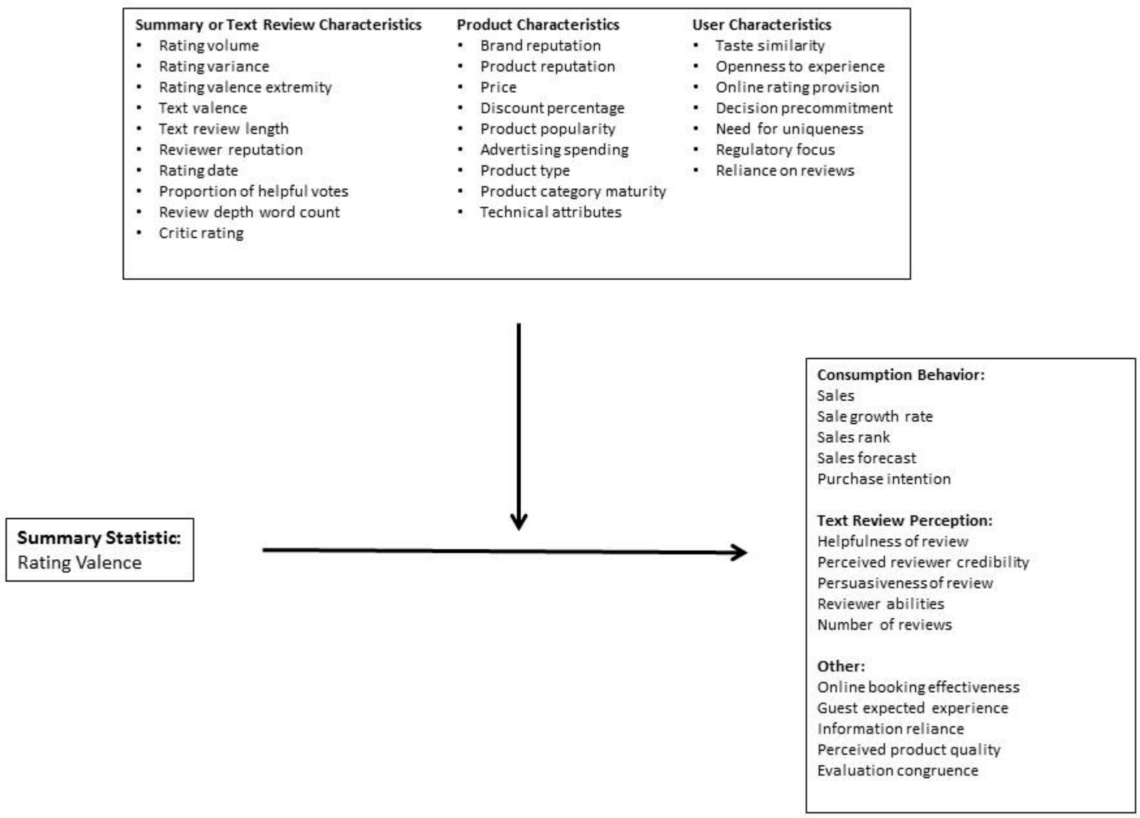


Table 2: Overview of Previous Quantitative OCR Literature by Article

				OCR Variables					
Article	**Study Objective**	**Data**	**Product**	**Valence**	**Volume**	**Variance**	**Interaction**	**Other Variables of Interest**	**DV**
Amblee and Bui (2011)	Effect of OCR on sales for low-cost digital products to remove effect of price on quality	Amazon	eBooks	No effect	Positive effect			Brand reputation; Product reputation	Sales
Chen et al., (2004)	Implications of OCR in the context of search costs for fit	Amazon	Books	No effect	Positive effect			Price; Discount percentage; Book popularity; Number of recommendations	Sales
Chen et al., (2006)	How social status impacts consumer responses	Amazon	Books	Positive effect	Positive effect			Proportion of helpful votes; Book popularity; Reviewer reputation; Spotlight Review	Sales
Chen et al., (2011)	Interactive effects of WOM and observational learning	Amazon; CNET	Camera	No effect	Positive effect			Observational Learning	Sales
Chevalier and Mayzlin (2006)	Effect of OCR on sales for Barned and Noble versus Amazon	Amazon; Barnes and Noble	Books	Positive effect	Positive effect	Positive effect		Length of review, Recency of rating	Sales; Book Rank
Chintagunta et al., (2010)	Measure the impact of OCR on box office performance of movies	Yahoo Movies	Movies	Positive effect	No effect	No effect	No effect	Advertising spending; Number of Theaters; Critic Scores; Days since release	Sales
Clemons et al., (2006)	Effect of review variance on beer sales using hyperdifferentiation	Ratebeer.Com	Beer	Positive effect	No effect	Positive effect	Significant		Sales growth rate
Clemons and Gao (2008)	Effect of review variance on online hotel reservations using hyperdifferentiation	TripAdvisor	Hotels	No effect	No effect	Positive effect		Absence/presence of strong positive/negative reviews	Sales; Online booking effectiveness: Guest expected experience
Cui et al., (2012)	Effects of online reviews on new product sales	Amazon	Consumer electronics; Video games	Positive effect	Positive effect			Product type (experience vs search)	Sales
Dellarocas et al., (2007)	Ability of eWOM to forecast box office sales	Yahoo Movies	Movies	Positive effect	Positive effect				Sales
Duan et al., (2008)	WOM leads to sales which leads to WOM (endogenous, positive feedback mechanism)	Yahoo Movies	Movies	No effect	Positive effect			Prerelease marketing costs; Number of theaters; Number of celebrities; Other movie characteristics	Forecasted Sales
Flanagin and Metzger (2013)	Investigate credibility of evaluations from user-generated content	Survey	Movies	Positive effect	Positive effect			Frequency of online information provision	Perceived credibility; Information reliance; Evaluation congruence; Behavioral intentions

				OCR Variables					
Article	**Study Objective**	**Data**	**Category**	**Valence**	**Volume**	**Variance**	**Interaction**	**Other Variables of Interest**	**DV**
He and Bond (2015)	Introduce moderators to explain effect of dispersion	Survey	6 various taste-similar or taste dissimilar products			Positive effect		Product domain (taste similarity); Review attribution (product vs reviewer); Openness to experience	Purchase intention; Product Evaluation
Ho-Dac et al., (2013)	Effect of OCR on brand strength across emerging and mature products categories	Amazon	Blu Ray and DVD players	Positive effect	Positive effect		Significant	Brand equity; Product category maturity	Sales
Khare et al. (2011)	Impact of volume on negative ratings	Survey	Movies	Positive effect	Positive effect		Significant	Message consensus; Decision precommitment; Need for uniqueness	Movie preference
Kostyra et al., (2016)	Effect of OCRs (and interaction) on product choice	Survey	eBook reader	Mixed effects	Positive effect	Mixed effects	Significant	Brand; Price; Technical attributes	Choice Probability
Moe and Trusov (2011)	Social dynamics (unrelated to objective assessment) observed in ratings and their effect on sales	National retailer's website	Bath; fragrance; beauty products	Positive effect	Positive effect	Positive effect	Significant	Product type (hedonic vs. utilitarian)	Sales; Rating behavior
Mudambi and Schuff (2010)	What makes an online review helpful to consumers?	Amazon	6 electronics products	Mixed effects				Product type (search vs. experience); Review depth word count	Helpfulness of review
Schlosser (2011)	Review persuasiveness using positive text reviews instead of presenting pros and cons	Yahoo Movies; Survey	Movies	Mixed effects				One vs two-sided written argument	Persuasiveness of review; Reviewer abilities
Sun (2012)	Effect of rating distribution and OCR factors interaction	Amazon; Barnes And Noble	Books	Positive effect	Positive effect	Negative effect	Significant	Price	Sales rank
Ye et al., (2009)	Effect of online consumer-generated reviews into online hotel booking services	Ctrip	Hotels	Positive effect		Negative effect		Price; Hotel star rating	Sales; Number of bookings
Zhang et al., (2010)	Persuasiveness of eWOM using regulatory focus theory	Experimental survey; Amazon	Photo software; anti-virus software	Positive effect			Significant	Regulatory (promotion vs. prevention)	Persuasiveness of review
Zhang et al., (2013)	Test effects of a search good with objective properties instead of experience goods which are subjective.	Amazon	Digital Camera	Positive effect	Positive effect			Price; Camera properties	Sales
Zhu and Zhang (2010)	How product and consumer characteristics moderate the influence on sales for video games	GameStop	Video games	Positive effect	Positive effect	Negative effect	Significant	User internet experience; Video game popularity; Reliance on reviews	Sales

In addition, because of the greater number of findings to report for positive valence and positive volume effects, we have further organized these respective sections by product, namely movies, books, electronics, and miscellaneous. Remaining sections which report negative or non-significant effects, including the variance section, will not be organized by the type of product given the lesser number of findings to report.

Valence

Below we will discuss the positive, negative and lack of effects of valence on consumption behavior for a variety of experience and search goods.

Positive Effects of Valence

Movies. Dellarocas et al. (2007) investigate if eWOM can aid in the forecasting of box office sales using movie reviews and statistics from Yahoo Movies. Using traditional WOM theories, they argue that individuals who live together share similar brand preferences. To explore this theory in a movie-goer context, the authors combine online review metrics with theater count and professional critic reviews to model more accurate revenue forecasts by geographic region. They show that the summary ratings are statistically significant as a predictor of sales forecasts and they further conclude that total box office revenues can be predicted from user reviews in the first week of a movie release (Dellarocas et al. 2007). Similarly, Chintagunta et al. (2010) also find national online review ratings to positively impact box office performance of movies, without any significant interactions to mention. The authors used daily box office ticket sales data for 148 movies released in the United States during a 16- month period, again collected from the Yahoo Movies website. In contrast with previous studies that have largely found that the main driver of box office performance is the volume of reviews, they find that it is the valence that seems to matter (Chintagunta et al. 2010). In addition, Flanagin and Metzger

(2013) argue that perception of movie reviews is based on an individual's experience with online information provision. It is implicit in user-generated content theories that a collective benefit will emerge from aggregated contributions (Flanagin and Metzger, 2013). From a random sample of over 1,000 adults with online-access, results indicate that an individual's movie ratings and their behavioral intentions are positively related; however, this relationship is greater for individuals who have more experience providing online reviews. The relationship between online rating experience and rating evaluations is understandable, as people are more likely to provide ratings and reviews may also be more likely to recognize and use OCR.

Books. Chen et al. (2006) study how social status impacts consumer responses to book reviews. Although their research is primarily concerned with the qualitative aspects of OCR, instead of the rating, they do test the interaction effects with summary ratings. Data is collected on book sales from Amazon and findings show a positive effect of valence on book sales. In another book review study, Chevalier and Mayzlin (2006) examine the effect of ratings on relative sales, specifically investigating if negative reviews on Barnes and Noble's and Amazon's websites will decrease sales more than increase of sales from positive reviews. They measure sales of approximately 5,000 randomly chosen books sold online between the two firms and find that sales improve when books have positive ratings, yet the results are non-significant for Barnes and Noble. They find an overall positive effect of valence on book sales, and especially notice that the negative impact of one-star reviews is larger than the positive impact of five-star reviews. Likewise, Sun (2012) investigate the psychological underpinnings of rating distributions on consumer evaluations for book sales drawing upon theories of product fit and match, specifically to test the effects of variance. The author finds that a higher summary rating on Amazon is positively related to book sales for Amazon but not for Barnes and Noble.

Additionally, the author finds that a product with a low summary rating and a higher variance communicates to potential buyers that well-matched consumers would favor the product.

Electronics. Ho-Dac et al. (2013) study the effect of OCR on brand strength across emerging and mature products categories of Blu-ray players on Amazon.com and results indicate that brand equity moderates the relationship between OCRs and sales. They rely on brand signaling literature to emphasize the product quality uncertainty that consumers face and how ratings and reviews may minimize this perceived risk (Shimp and Bearden, 1982). Positive (negative) OCRs increase (decrease) the sales of models of weak brands but do not have a significant effect on the sales of models of strong brands. Also, Zhang et al. (2013) test effects of a search good (digital camera) using sales data from Amazon. They reiterate the notion that experience good evaluations are highly subjective in nature, whereas search goods are evaluated by their objective properties. The literature that examines OCR influence on experience goods contains mixed effects for valence and volume. The authors believe that search goods may be better suited to online evaluations, because reviews of search goods address objective aspects of the product. Their results reveal that not all online reviews of an experience good are important, yet change in price and valence are significantly associated with future sales for a search good. In another study, Zhang et al. (2010) apply regulatory focus theory to test the effects of OCR variables on anti-virus software sales using Amazon data. They show that for products associated with promotion consumption goals, consumers rate positive reviews as more persuasive than negative ones. Conversely, consumers rate negative reviews as more persuasive than positive ones (negativity bias) for products associated with prevention consumption goals (Zhang et al. 2010). Similarly, Cui et al. (2012) investigate the effects of online reviews on new product sales of consumer electronics and video games using WOM theory. In their study, data was collected using Amazon sales figures and the authors find a positive effect of valence on video game sales, such that sales increase as the summary rating increases. Outside of the Amazon context, Zhu and Zhang (2010) seek to answer how product and consumer characteristics moderate the influence of OCR on sales for video games using data found on GameStop.com. They base their framework on the psychological choice model, which states that consumer's search effort is influenced by their product knowledge. Their primary findings show support for a more positive effect of valence on sales when consumer internet experience is greater.

*Miscellaneous***.** Clemons et al. (2006) examine the effects of resonance marketing on beer sales using data from ratebeer.com. Resonance marketing occurs when products are developed to produce the strongest favorable responses among a smaller segment of consumers (niche), where only the most informed consumers find what they are looking to purchase. The authors find a positive effect of valence on beer sales growth rates, but only for the top quartile of ratings. In a different study, Moe and Trusov (2011) use data from a national retailer's website to explore social dynamics observed in ratings and their effect on sales and subsequent reviews of bath, fragrance, and beauty products. Specifically, their analysis shows that positive ratings result in higher products sales and subsequently more positive reviews. Furthermore, an increase in the volume of ratings can offset the negative effect of a decrease in valence on sales. Finally, Ye et al. (2009) analyze the effect of online consumer-generated reviews using online hotel booking services with hotel review data from Ctrip, the largest travel website in China. They state that reviews for experience goods are especially important as information regarding their quality is often unknown prior to purchase. Results show that positive online reviews can significantly increase the number of hotel bookings. The results further suggest that a 10% improvement in summary ratings can increase sales by over 4%, although room rates had a negative impact on the number of online bookings.

Studies Reporting a Lack of Effects of Valence

Duan et al. (2008) appraise how eWOM leads to sales which subsequently lead back to more eWOM in a positive feedback mechanism. They are primarily interested in forecasting sales and study these effects using movie data of 71 movies found on the Yahoo Movies website. Results show that valence has no significant impact on movies' box office revenues, indicating that online user reviews have little persuasive effect on consumer purchase decisions. This finding contradicts the positive effects of valence on movie sales found by Chintagunta et al. (2010). In addition, Clemons and Gao (2008) present a study of online hotel reservations using 13,728 reviews on TripAdvisor.com to show that overall valence does not appear to be correlated with online booking effectiveness. Using camera sales and reviews from Amazon and CNET, Chen et al. (2011) draw upon interactive effects of WOM and observational learning concepts. Observational learning states that a consumer’s purchase decision is influenced by the actions of others, such as the reported sales percentages of consumers after viewing a product online. They find no clear evidence that the impact of others’ purchase actions will increase when consumer ratings increase.

Chen, Wu and Yoon (2004) study the implications of the OCR recommendation systems in the theoretical context of search costs and product fit using book sales data on Amazon. Interestingly, consumer book ratings are not found to be related to sales. The authors believe due to different consumer tastes, shoppers may get a book they like regardless of other’s ratings and reviews. Also, since most of the books receive relatively high ratings, consumers may not find these ratings helpful. In another context, Amblee and Bui (2011) investigate the influence of

OCR on sales for low-cost eBook products to remove the effect of price on quality. They use concepts of social influence and its effect on social commerce to explain potential effects on eBook sales. Social influence is the process by which individuals make changes to their attitudes and behaviors to align with other individuals or groups. The Amazon data that they used contained a 4 out of 5 rating for over 90 percent of all reviews and valence was not found to be statistically significant. Possibly, a lack of variability may account for the non-significant predictive power of the summary rating.

Across several product category types, the effects of valence on dependent measures, like sales, appear to mostly positive. In general, as summary ratings increase so do consumer perceptions of the product. Although some studies show a lack of effect of valence, there are no reported negative effects of valence.

Volume

Positive Effects of Volume

Movies. Dellarocas et al. (2007) analyze secondary data from Yahoo Movies and indicate a positive effect of volume on box office sales for a movie with positive valence. Interestingly, they also show that volume of online reviews can be used as a proxy of sales. More results using movie reviews and sales information on Yahoo Movies find that box office sales are significantly influenced by the volume of online postings, suggesting the importance of what can be called an ‘awareness effect’ which relates to existing WOM theories regarding an individual’s propensity to seek and follow the opinions of others (Duan et al. 2008).

Movie review data measured from online surveys also show significant positive effects of volume, like results from secondary data. Flanagin and Metzger (2013) use an online survey regarding movie reviews to show that the OCR volume is positively associated with perceived credibility of the reviews and confidence in the accuracy of the ratings and text reviews. Although these variables are not directly tied to sales, they can arguably influence subsequent purchases. Khare et al. (2011) also use an online survey about movie reviews to show several interactive effects with unique user characteristics. Results of their studies indicate that an increase in volume will likewise improve an individual’s preference for a product that is positively rated. On the other hand, an increase in volume will reduce consumer preferences for a product that is negatively rated. Finally, they find that consumers with a greater need for uniqueness in their consumption are less susceptible to follow the opinions of others.

Books. Amblee and Bui (2011) find that the regression between sales and the volume of customer reviews for Amazon eBooks is statistically significant. Holding valence constant, the total volume of reviews posted for an Amazon Short can explain 15.9 percent of the variance in sales. Authors of a different study also discover that the number of reviews a book has on Amazon is also found to be positively related to sales (Chen et al. 2004). People seem more likely to discuss a book that is currently popular in online discussions, thus providing sales momentum in the market. However, greater number of reviews may be influenced by greater number of sales, so it is difficult to infer a causal relationship between volume and sales (Chen et al. 2004). In addition, findings of Chen et al. (2006) suggests that volume of reviews is positively related to book sales on Amazon, but the marginal impact of an additional review declines with the number of reviews. Chevalier and Mayzlin (2006) similarly use secondary data to research the effect of OCR on book sales, relative to Amazon and Barnes and Noble, and show how volume improves sales for Amazon alone. Additionally, Sun’s (2012) study of book reviews on Amazon found that when the summary rating is positive, a higher number of reviews led to higher sales, as we have seen in other studies. Lastly, a single experimental survey (Kostyra et al.

2016) shows similar positive effects of volume as seen in the previous Amazon datasets. They find an interactive effect between volume and valance, where the positive effect of volume on choice probability for an eBook reader is seen at high levels of valence.

Electronics. Chen et al. (2011) use Amazon digital camera data and find that volume is a significant predictor of sales for consumer electronics and video games on Amazon. Specifically, valence has a stronger impact than volume for search goods, which consumers can evaluate by specific attributes before purchase; however, for experience goods this effect is reversed as experience goods require more feeling or experience to properly evaluate. Ho-Dac et al. (2013) find a significant interaction between volume, valence and brand strength, for Blu-ray and DVD player reviews and sales on Amazon. Results indicate that more sales lead to a larger volume of positive (but not negative) OCRs, which then lead to greater sales again, but only for weak brands. This loop does not exist for strong brands, because they seemingly do not benefit as much from positive reviews. Again, using Amazon data, Zhang et al. (2013) also show that the number of online reviews is positively related to digital camera sales. Zhu and Zhang (2010) consider the video game market using data from GameStop.com and find positive effects of volume on sales. However, less popular games seem to benefit more from this volume increase than do popular games, given the greater need to use OCRs for unfamiliar brands to obtain quality information to reduce purchase risk (Zhu and Zhang, 2010).

Miscellaneous. Moe and Trusov (2011) analyze bath, fragrance and beauty product reviews from a national retailer's website to study how people make reviews based on previous reviews. Results indicate that volume has a positive effect on sales and yet a negative effect on consumer's likelihood to provide additional ratings.

Studies Reporting a Lack of Effects of Volume

Chintagunta et al. (2010) studied reviews from Yahoo Movies and found that it is the valence that seems to matter for box office revenues and not the volume, which contrasts with findings of previous studies. Their results were found holding valence constant and using local market-level data only. Yet, when they aggregate the data across markets (national-level data) they find positive effects of volume. They use these conflicting findings to show that results depend on which method of aggregation is used and discuss how biases in aggregated data may be overcome. Clemons et al. (2006) examined the craft beer industry and found a positive effect of valence and variance on sale growth but not volume. The authors believe that it is more important to have a few loyal customers instead of a larger number of impartial customers. Similar results are found in the hotel market, as Clemons and Gao (2008) point to their results, from TripAdvisor data, which identify that it is not the number of reviews that serves as a proxy for online marketability but valence and variance.

In general, the effects of volume appear to be mostly positive, especially for products with positive summary ratings. Few studies show a lack of effect of volume and there are no reported negative effects of volume.

Variance

Compared to the larger stream of experimental and secondary research on valence and volume, little work has focused on the variance of OCR ratings and reviews (He and Bond (2015). We will now discuss the positive, negative, and lack of effects found in the literature for the variance OCR variable.

Positive Effects of Variance

Clemons and Gao (2008) use TripAdvisor.com to examine effects of variance on hotel bookings and sales growth. Using the concept of resonance marketing, they suggest that sales of a firm can benefit from a smaller portion of positive evaluations of a product, despite mixed ratings, and that variance is positively associated with sales growth. Although a firm may have mixed ratings and high variance, the product offering of the firm may resonate with a certain segment of consumers. For example, some consumers may perceive Whole Foods Market as an overpriced grocer, while others may perceive it as a worthwhile organic alternative to unhealthy foods. Clemons et al. (2006) also draw upon resonance marketing and find a positive effect of variance in their analysis of craft beer data on Ratebeer.com. Using an experimental survey for a variety of taste-similar and taste-dissimilar products, He and Bond (2015) find positive effects of rating variance on product-related judgments and choice. Taste similarity is the extent to which evaluations in a product domain are expected to differ among consumers. For example, a lamp may have higher levels of product taste-similarity, whereas a painting may have higher levels of taste-dissimilarity. In this case, participants were more likely to choose a product with a high variance of ratings when the product domain was characterized by dissimilar tastes, because it is plausible for such a product to show rating variance. In a concluding study, the authors introduced a consumer characteristic (openness to experience) and show participants high (low) in openness responded favorably (unfavorably) to variance when the variance could be attributed to the individual reviewers and not the product. Moe and Trusov (2011) similarly find a positive effect of variance on sales, but a negative effect of variance on both extreme (e.g., 1 or 5 out of 5) rating helpfulness and future rating behavior. When consumers view ratings with greater variance, they may be less influenced by extremely negative or extremely positive ratings. Sun

(2012) finds that variance is positively related with book sales on Amazon when the summary rating is low, and hurts relative sales when the summary rating is high. Drawing upon theories of product fit, the author demonstrates how a product with a low summary rating and a high rating variance may act as a signal of “fit” for well-matched consumers. On the other hand, for a product with a high summary rating, a high variance of ratings may reduce demand.

Negative Effects of Variance

In a study using book sales rank from Amazon and Barnes and Nobles, Sun (2012) finds that a higher standard deviation of ratings on Amazon improves the book’s relative sales on Amazon when the summary rating is low, and hurts its relative sales when the summary rating is high. More specifically, a higher standard deviation of Amazon ratings increases the book’s relative sales when the average Amazon rating is lower than 4.1 stars. On the other hand, the author finds that a higher summary rating on Amazon increases the book’s relative sales when the standard deviation is lower than 1.6 stars (Sun, 2012). Ye et al. (2009) also show a negative effect of variance results for hotel sales, using Ctrip data. Specifically, results suggest that a 10% increase in review variance can decrease sales by 2.8%. For video game sales, Zhu and Zhang (2010) also find a negative effect of variance, using data from GameStop.com. This negative effect of variance is more prominent for less popular video games.

Studies Reporting a Lack of Effects of Variance

Despite other literature that has found positive and negative effects of variance on consumption behaviors, Chintagunta et al. (2010) find that variance has no effect on future box office performance. The use movie reviews and sales data on Yahoo Movies, specifically analyzing local markets and not national-level aggregate data.

Overall, valence and volume are shown to have mostly positive effects on product sales and other measures of interest. Fewer studies report a lack of effect for valence and volume, without any reported negative effects for these OCR variables. In addition, the positive effect of volume is greater when summary ratings are positive. Interestingly, reported effects of variance were also mostly positive, showing that purchase behavior increases as the divide between positive and negative ratings increases, but only when the average summary rating was less positive. As would be expected, some negative effects of variance were found, like in the case of an unpopular product. Lack of effects was reported for each of the three OCR variables, but only make-up a small portion of the research findings. Although WOM is not a new research domain, eWOM is. It is plain to see that although there is much support for positive, negative and no effects of all three OCR factors (valence, volume and variance) there appear to be mixed results.

CHAPTER 3: RESEARCH QUESTIONS

Scale length research is inconclusive and supports the use of scale lengths anywhere from 2- to 25-point to accurately capture a respondent's evaluation. Furthermore, it remains to be seen how different scale lengths affect consumers who are evaluating OCR. It is evident that OCR variables are a popular subject of research, given the numerous studies on this topic and OCRs pervasive use and influence in the market. As we have already discussed in our review of the OCR literature, researchers have explored OCR variables in many contexts, using both secondary and primary data. Researchers have also analyzed the effects of OCR variables for both 5- and 10-point rating scale lengths. However, there appears to be a gap in testing the effects of OCR summary ratings on pre-purchase consumer perceptions across different lengths of a product rating scale. In this chapter, we will first discuss the mixed findings in scale length research and how this gives rise to our primary research question regarding scale length comparison. Next, we will discuss different scale lengths that are primarily used in the OCR literature and the market, and why an analysis of the effects of scale length on consumer product preferences is important to investigate.

Effects of Scale Length

More than 100 years of research has studied the effect of scale length on a respondent's ability to provide the most precise evaluation. Although there have been many studies on optimal scale length over the last century, there still appears to be mixed results as to which scale length is most efficient at capturing individual assessment. Garner and Hake (1951) state that the amount of information conveyed by a scale has been found to increase with an increase in the number of response categories (scale points, like 5- or 10-point). Additionally, a meta-analysis found that the reliability of a scale increased with an increase in the number of response categories (Churchill and Peter, 1984). However, other research claims that an increase in number of scale categories would require greater evaluative effort and thus may be too difficult for respondents to properly record an accurate response (Park and Lessig, 1981). Similarly, more current research argues for the use of shorter scales than the 5- or 7-point scales often used in research, because they may be easier to administer and easier for consumers to complete (Viswanathan et al. 2004).

Komorita and Graham (1965) point out that a scale with too few categories does not allow for sufficient discrimination between scale categories whereas a scale with too many categories may be beyond the consumers' ability to discriminate. The optimal number of response categories could vary anywhere from 3 to 25, depending on individual preferences (Viswanathan et al. 2004). However, this research has examined the role of scale length in the context of a consumer providing a rating using a scale but not specifically in the context where a consumer is evaluating ratings provided by other consumers. Nonetheless, intuitively, it seems reasonable to expect that if scale length can affect individuals' provision of ratings, they might also influence their evaluations of ratings provided by others.

OCR in the Literature

Much of the literature that investigates OCRs using a 10-point rating scale pertain to experience goods, as much of the online market employs 10-point rating scales for experience products (e.g., Yahoo Movies, iMdb, Priceline, and Rotten Tomatoes). Conversely, Mudambi and Schuff, (2010) use secondary data from Amazon and therefore use only a 5-point rating scale. Early OCR literature (2000-2010) and even more recent OCR literature (2011-current) gives some attention to the 10-point rating scale. Authors like Dellaracos et al. (2007), Duan et al (2008), and Chintagunta et al. (2010) are among several whose OCR research utilizes a 10-point rating scale. Furthermore, authors have studied the effects of additional factors like product type (Ye et al. 2009), price (Mudambi and Schuff, 2010), regulatory focus (Zhang et al. 2010), and brand equity (Ho-Dac et al. 2013), to name a few. Although this research better explains the interactive relationships among OCR variables (valence, volume, and variance) it lacks any testing or discussion on the effects of scale length on product evaluation using consumer ratings.

OCR in the Digital Marketplace

After performing a search of rating and review sites within the online consumer market, it was obvious that most OCR platforms utilize a 5-point star rating scale as the preferred scale length to capture and showcase consumption experiences (see Table 1.1). In general, rating scales allow consumers to both provide and review product evaluations regarding their online shopping experience. Although our search was not exhaustive, it was an extensive list of approximately 90 sites, including both eCommerce (e.g., Overstock) and crowd-sourced review sites (e.g., Yelp). In addition, it comprised both traditional brick-and-mortar retailers who also offer products online (e.g., Costco) and purely online retailers (e.g., Amazon), all offering a variety of product categories.

Most organizations in the online market use a 5-point star rating scale for online product evaluations. The remaining organizations predominately used a 10-point star rating scale. Arguably, ecommerce giants like Amazon may have set the 5-point rating scale precedent which may explain why most other sites elect to conform to the standard set by the online market giant. Although this may lead some to discount the 10-point scale as a viable evaluative tool, its effect on online purchase behavior is worth examining.

Rating Scale Length

OCR research has contributed volumes of work over the past decade, but has only explored the effects of summary ratings for a single scale length (usually 5- or 10-point). Consequently, there appears to be lacking a comparison in the effects of rating scale length on consumer product perceptions. This leads us to our primary research question:

1. Does scale length (5- vs 10-point) effect consumers' perception of product quality and purchase intention when assessing a product summary rating?

We are not interested in the effect that scale length has on how consumers decide to provide a rating for a product they have already purchased but how potential customers evaluate the ratings already provided by previous consumers. Some may argue that a comparison of scale lengths is unimportant if the summary rating between two scale lengths is equivalent (e.g., 4 out of 5 vs. 8 out of 10 are equivalent proportions). However, many online product review sites will only provide a visual scale and not specify the summary rating using numerical values. For example, a product evaluation may include a 5-star (10-star) rating scale, with 4 out of 5 (8 out of 10) stars colored yellow, representing an 80% rating score, without providing a numerical value next to the scale to indicate the rating score. For our first study, we start with the assumption that consumers come to learn the summary rating by viewing a displayed visual scale and not a displayed numerical value. Per Table 1.1, this is a valid market assumption. Later studies will explore the effects of rating scale length when the rating percentage is provided as well.

If a product was rated using a 5- and 10-point scale length, and the summary ratings were proportionately equivalent (e.g., 4 out of 5 and 8 out of 10), then we would assume that consumer perceptions of this product would not differ across scale length. However, prior mathematical and psychological research points to the use of heuristics used to evaluate proportions. Given that a visual scale (i.e., 5 or 10 stars on display) may not include a numerical rating beside it, consumers can interpret the summary rating in terms of a proportion involving numerals (e.g., “this product received 4 out of 5 stars”) or interpret it purely geometrically based on the distance of the right most shaded star from the left or right-endpoint in relation to the total measured length of the scale. Thus, the visual processing relies on a cognitive representation that is free from numerals. If consumers make any mistake interpreting the scale proportion, it may lead to either underestimating or overestimating the actual summary rating. This leads us to our second research question:

2. If perceptions of product quality and purchase intentions vary by rating scale length, then what process influences these evaluative differences?

If consumers view a 4 out of 5 rating using a 5-point scale then we will also examine an 8 out of 10 rating using a 10-point scale, etc. Those who recognize these proportions as equivalencies may fail to understand any reason in devoting a study to comparing their differences. Yet, we live in a world where consumers prefer heuristic processing over numeral-based processing, for a variety of reasons (Petty et al. 1997). These conditions may lead to a disparate perception of quality between products of the same summary rating percentage, but different scale lengths. In the case of OCR ratings, as consumers compare an equivalent summary rating from different scale lengths (e.g., 4 out of 5 vs. 8 out of 10), heuristic-based processing of these proportions may lead to different perceptions of quality for a specific product.

Boundary conditions may exist where the core effect of scale length becomes absent because respondents no longer rely on the summary rating; when more concrete or informative information is presented alongside the summary rating scale or some other factor makes reliance on the scale risky or not useful. For example, a summary rating percentage (e.g., "80%" when the rating is 4 out of 5 or 8 out of 10) or volume of individual raters (e.g., "15 individuals") could be displayed next to the 5- and 10-point summary rating scale. Our third and final research question is:

3. Will the presence of a rating percentage or varying levels of rating volume moderate the effects of scale length on perceptual outcomes?

When such information is present, consumers might not rely on the scale because more concrete information that is not subject to interpretation is present (in case of the percentage-based information) or the volume of raters is so low as to make the scale ineffective for quality judgment, regardless of the scale's length. The following chapter comprises the concepts and hypotheses regarding our research questions.

CHAPTER 4: CONCEPTUAL BACKGROUND AND KEY HYPOTHESES

In this section, we discuss the various theoretical perspectives that can potentially explain how people might perceive product summary ratings and offer relevant hypotheses for our research. It is possible that consumers may consider converting an OCR proportion into a rating percentage (e.g., 4 out of 5 is 80%) to gain a better sense of the summary rating. If this is the case, we argue that visual heuristic processing (a shortcut strategy to provide an estimate) will allow for a less effortful evaluation of the summary rating. However, even if consumers are not willing or able to convert the summary rating proportion into a percentage, visual heuristic processing will still be a factor in their assessment of the rating. We propose that visual heuristic processing will be manifest in terms or endpoint anchoring using the visual OCR summary rating scale. Our argumentation for the probable effects of scale length will lead to competing hypotheses.

Computational Ease

Online consumer ratings allow buyers to provide a personal product evaluation in which potential buyers can base their judgments on those evaluations. Given that a product receives a 4 out of 5 rating, for example, consumers may process this rating in several ways. Clearly, 4 out of 5 is a mathematical proportion, or fraction, that can be converted to a percentage (i.e., 80%). This percentage can be used by consumers as a signal of product quality to then make purchase decisions. However, cognitive psychology literature would argue that adults find fractions difficult to process (Bonato et al. 2007). Adults may find that calculating the rating percentage is difficult, or they simply lack the ability or motivation to compute, and therefore may not perceive a 4 out of 5-star rating to be 80% (assuming the rating percentage is not displayed). Possibly, consumers may calculate, or estimate, 4 out of 5-stars to be lower or higher than the true percentage, which may influence their evaluation of the product. This idea of computational ease is worth investigating in a quantitative OCR context. Computational ease literature further explains that when consumers are unable or unwilling to use numeral-based processing, they may evaluate numerical proportions (like online ratings) more heuristically (Bonato et al. 2007).

As the old joke goes, five out of four people have trouble with fractions (Ischebeck et al. 2009). The authors continue to explain that fractions may be difficult for children and adults to understand because they are represented differently from other numbers or quantities in the brain. That is to say, it is much easier to think in terms of discrete numbers than in terms of fractions, proportions or rates (Bonato et al. 2007). It may be assumed that the difficulty in mastering fractions is specific to children and would not be an issue for educated adults. However, Gigerenzer & Hoffrage (1999) suggest that fractions and proportions are hard to understand even for adults, as "humans seem developmentally and evolutionarily prepared to handle natural frequencies but not proportions."

Individuals regularly come across fractions in daily life, such as part-whole relations and measurements (e.g., half an hour), proportions expressed as percentages (e.g., 15%), and chances (e.g., 1:4) (Ischebeck et al. 2009). A failure to understand the basic concept of fractions may cause difficulties in everything from cooking and time-management, to even qualifying for employment. Given the prevalence of eWOM summary ratings (e.g., 4/5 or even 4 out of 5 stars) in today's digital marketplace, it can even be argued that a failure to understand fractions and proportions may create difficulties in interpreting online summary ratings.

Bonato et al. (2007) show in a study that when fractions are compared side-by-side, adults often compare either the numerator or denominator to make judgments of the magnitude of the fractions. The authors argue that representing the meaning of a fraction in this way implies that the real value of the fraction is not readily accessible to these individuals. This provides initial support for the proposition that adults do not prefer numeral-based processing and seek for some sort of simpler approach to solving the problem, possibly a heuristic.

Heuristic Processing

We refer to the computational ease to point toward the tendency of heuristic use when faced with numeral-based processing. Star rating scales are fractions that highlight the proportion of a product’s summary rating to the total scale length. Bagchi and Davis (2012) state that when individuals have real or perceived difficulty performing numeral-based processing they use heuristics to make inferences. In a study conducted by Bonato et al. (2007) the authors show that even skilled participants prefer to take recourse to heuristics and do not access the real number that a fraction represents.

Pricing literature also discusses the use of heuristics to avoid numeral-based processing. For example, if the difference between the regular and sale prices is not specified in either absolute dollar or percentage terms, consumers frequently employ mental heuristics to avoid the effort of calculating the difference (Coulter and Coulter, 2007). The literature does not specify if individuals tend to underestimate or overestimate in the face of numeral-based difficulty, but it simply points to heuristics.

Building on heuristic processing, the anchoring effect explores how individuals use an initial stimulus to make subsequent evaluative judgments. In a shopping context, for example, consumers may be influenced by an initial value that acts as an anchor, and any fluctuation from that starting value will then be judged relatively. Perhaps, online consumers anchor toward an upper or left-endpoint on the visual rating scale which influences their assessment of the summary rating. The possibility of consumers choosing to anchor on some numerical value or visual point on the rating scale is the basis of our study.

Anchoring Effects

Anchoring was proposed by Tversky and Kahneman (1975) as a heuristic used in judgment to simplify calculations. Anchoring research suggests that individuals regularly anchor on the first bit of information presented, form initial judgments, and then fail to update those judgments to account for subsequent information (Bagchi and Davis, 2012). This means that although anchoring may ease the burden of numeral-based processing, the estimated result may be numerically inaccurate.

Anchoring effects have been actively studied since the 1940's, in the context of physical magnitude differences (e.g., Heintz, 1950; Helson, 1948;) and even the effects of numbers on communication and persuasion (e.g., Hovland, Janis and Kelley, 1953). More recently, we have seen anchoring effects applied to the marketing literature in a pricing context (e.g., Janiszewski and Lichtenstein, 1999; Bagchi and Davis, 2012). Epley and Gilovich (2010) applaud the efforts of authors who applied anchoring theories to new domains and contend that it is important to study the effects of anchors encountered in everyday life. We propose extending the findings of anchoring as a heuristic to the issue of people's judgment based on ratings, particularly in relation to the use of upper and left-endpoints anchors on OCR rating scales.

Rating scales include a range of values, from a left to a right-endpoint (e.g., 1 to 5 or 1 to 10). In general, the range of a set of values determines the perceived value of any one stimulus in the range (Janiszewski and Litchenstein, 1999). For example, a speed of 60mph may be perceived as fast when the range stretches from 20mph to 70mph, but slow when the range is from 60mph to 70mph (Janiszewski and Lichtenstein, 1999). This study uses the same speed

(60mph) to examine its relative location within two different ranges, but not the same speed-to-range proportion in different ranges-which we seek to investigate, per our first research question previously discussed. Janiszewski and Lichtenstein (1999) also believe that the end anchors govern the major properties of the judgment reference scale which the individual adopts in the rating of his or her attitude.

Endpoints as Anchors

Sherif and Hovland (1961) also argue that it is the end values of the range that usually acquire an anchoring role. Thus, endpoints seem to exert influence over individual's judgments regarding a value within a range (Ostrom and Upshaw, 1968). Like the above example of a speed range, Ostrom and Upshaw (1968) show comparable results in a range of evoked prices. The results of their price range study show that when the upper bound of the range of evoked prices increased, the perceptions of a certain market price become more favorable, and when the lower bound of the range of evoked prices is decreased, perceptions of the same market price become less favorable. In other words, the attractiveness of a market price changed as the price range changed, even though there was no change in the reported reference price (Janiszewski and Lichtenstein, 1999).

There is little doubt that numerical anchors influence subsequent judgments (Epley and Gilovich 2010). Tversky and Kahneman (1974) illustrate how individuals often anchor on the first piece of information provided. They show such an effect as study participants give very different estimates of 8! depending on presentation order (1 x 2… vs. 8 x 7…). The descending sequence produced larger estimates, suggesting that individuals focused on the first piece of information as an anchor to make subsequent inferences (Epley and Gilovich, 2010). Tversky and Kahneman (1974) state that people form estimates by starting from an initial value which is adjusted to yield the final answer, a phenomenon called anchoring and adjustment. Whatever the source of the initial value, different starting points yield different estimates, which are biased towards the initial anchoring value.

If consumers are not provided a starting point anchor in a manipulated experimental condition, consumers may anchor on the right-endpoint or left-endpoint of the rating scale. Given that OCR rating scales signify product quality and consumer satisfaction, it would seem logical to predict that consumers will anchor on the right-endpoint of the rating scale (i.e., 5 or 10) which indicates more quality. On the other hand, the first value that consumers may see is the left-endpoint of the rating scale (i.e., 1), especially in a society inclined to process words and numerical scales in left-to-right fashion.

Right-Endpoint Anchoring

Coulter and Coulter (2007) describe a heuristic which involves comparing the numerical digits of two prices from right-to-left. In the results of their study, the authors show that if the left digits are the same, then more attention is focused on the right digits in the price comparison process. For example, when $23 is compared to $22 consumers will anchor on the right digit since the left digit is identical. Although online consumers who review a single product webpage are not comparing two-digit prices side-by-side, we use this pricing example as an indication that there is some empirical evidence to support that consumers may be prompted to anchor on the most-right value. Applying this directional preference has implications in OCR rating scales, as anchoring on the most-right value in a scale equates to right-endpoint anchoring. However, a simpler explanation for right-endpoint anchoring exists: since rating scales signify product quality, it would seem logical that consumers would anchor on the right-endpoint of a rating scale because of their desire to purchase a quality product. For example, a rating of 4 out of 5 (as an example of an 80% rating on a 5-point scale) may be more difficult to process numerically than the proportion 8 out of 10 (as an example of an 80% rating on a 10-point scale). If a proportion is difficult to compute, humans will possibly rely on a heuristic, and seeing that 4 is closer to 5 than 8 is to 10, they may overestimate the 4 out of 5 rating. This is a likely outcome of anchoring on the right-endpoint of the rating scale. However, much of the anchoring research finds that consumers anchor on the first value that they see. As such, we will also test if the first value that consumers see is in fact the lower scale endpoint (i.e., 1), given a natural inclination for individuals to engage in left-to-right processing.

Anchoring on the right-endpoint will lead to a comparison between the right-endpoint and the summary rating value. For example, if the rating is 4 out of 5 stars then the visual and magnitude difference between the right-endpoint and the rating (5-4=1) will appear less than that of the equivalent rating of 8 out of 10 stars (10-8=2). The smaller distance between the right-endpoint and the rating for the 5-point scale may lead to greater perceptions of product quality and purchase intention, compared to the 10-point scale. This leads us to our first, of two, competing hypothesis:

H1: Perceived product quality and purchase intention will be higher when the star rating is on a 5-point scale compared to a 10-point scale.

Left-Endpoint Anchoring

Fias and Fischer (2005) demonstrate that spatial and numerical processing are intimately connected. Furthermore, they reported this numerical processing was predominantly spatially oriented from left-to-right, in increasing order. For example, many cultures process numbers in left-to-right increments. Further evidence of left-to-right numerical processing by individuals can be found in the SNARC effect. The SNARC effect occurs as participants favor pressing a button located on the left of a keyboard for smaller values and a button located on the right of a keyboard for larger values (Fias and Fischer, 2005). The SNARC effect is much tested and is explained in this section of our paper to demonstrate that numerical magnitudes are spatially oriented in most people.

Consumers expect to see a number series in increasing order, from left-to-right (Biswas et al. 2013). Given this, consumers may extract judgments about the quality of a product by comparing the visual distance from the summary rating (e.g., 4 out of 5 stars) to the left-endpoint of the scale (i.e., 1). Since stars are often used as a proxy for a summary rating scale, we may also assume that consumers might naturally anchor on the left-endpoint as they evaluate the numerical differences using left-to-right processing.

OCR summary ratings are typically displayed as 5 horizontal stars. The summary rating is indicated by shading the appropriate proportion of stars, leaving the remainder blank (white). It is in this setting that we propose and explain the left-endpoint anchoring effect in OCR. In the case of using the left-endpoint as a rating anchor, a rating of 8 is farther from 1 than a rating of 4 is from 1, despite their proportionate equality. According to range theory as applied to pricing, consumers use the upper and lower bounds of a range to evaluate given its relative location within that range. In pricing, consumers would restrict the upper bound (more expensive) more than the lower bound (less expensive), but the opposite may be true for online consumers using a star display as a quality signal. For the online consumer, the rating bounds may be either the upper or left-endpoints of the scale. Although the ratings across the two scales are equivalent (e.g., 4 out of 5 and 8 out of 10), the visual dimensions of the scales and the distances between the summary rating and the scale endpoints vary.

Anchoring on the left-endpoint may lead to a comparison between the left-endpoint and the rating. For example, if the summary rating value is again 4 out of 5 stars then the visual and magnitude difference between the left-endpoint and the summary rating (4-1=3) will appear less than that of the equivalent proportion of 8 out of 10 stars for the 10-point scale (8-1=7). The smaller distance from the left-endpoint and the summary rating may result in weaker perceptions of product quality and purchase intention for the 5-point scale. This leads us to the second part of our competing hypothesis for Study One:

H1(ALT): Perceived product quality and purchase intention will be higher when the star rating is on a 10-point scale compared to a 5-point scale.

Based on our rationale and the literature supporting endpoint anchoring regarding the evaluation of OCR rating scales, it is also reasonable to expect that consumers may favor either a 5- or a 10-point scale, depending on which endpoint becomes their anchor. In Study One, we test the competing hypotheses regarding the effect of rating scale length on perceived product quality and purchase intention.

CHAPTER 5: STUDY ONE

Study One tested hypotheses H1 and H1(ALT) to explore the potential differences in consumer's evaluation of a product using OCR rating scales of different lengths (5-point and 10-point). Significant differences between the rating scale lengths may provide preliminary support for one of the competing hypotheses regarding either left or right-endpoint anchoring. In this chapter, details of the methodology and results of data analysis will be discussed.

Methodology

Experimental Design

A 2 (Rating scale length: 5-point; 10-point) x 2 (Rating percentage: 70%; 80%) between-subjects design was used for Study One. Summary ratings were manipulated using two rating percentage levels (70% for 3.5 out of 5 and 7 out of 10 summary ratings; 80% for 4 out of 5 and 8 out of 10 summary ratings) to provide more robustness to the study. An 80% rating percentage in this context, for example, does not imply that the percentage is displayed as a value next to the rating scale, only that respondents are viewing summary ratings that are either 4 out of 5 or 8 out of 10. The volume of ratings was held constant across conditions at 1,394 customer ratings, and was displayed alongside the star summary rating scale. The conditions related to the study were manipulated through an image designed to appear like an Amazon product webpage that potential buyers would view in an online purchase setting. The layout and details of the image and information for the product (electronic tablet) are consistent with a typical product webpage found on Amazon. In addition, the rating symbol (star), its color, size and position are congruent with the ratings scales used by Amazon, and found elsewhere in the market (See Appendix A).

Sample and Procedure

One hundred and eighty-five undergraduate business students (Female = 49%, M_{Age} = 23) from a large Midwestern university participated in the experiment. After the participants were randomly sorted to one of four conditions, they responded to some items that intended to measure their perceived product category. Next, each participant responded to questions related to purchase intentions and perceived product quality. In addition to the dependent variables of interest, we also measured online shopping experience (Zhu and Zhang, 2010) and brand familiarity (Ho-Dac et al. 2013) as additional variables that are used in the OCR literature.

Measures

All variables were measured using 9-point Likert scales, unless otherwise noted. *Purchase intention* and *perceived product quality* are the dependent variables of interest. *Purchase intention* (α = .91) was measured using three items adapted from Keller et al. (1997): “Imagine you were planning to buy a tablet. How likely would you be to buy to the isoTech tablet?” (1-not likely at all; 9-very likely), “How probable is it that you would consider the purchase of this product?” (1-not probable; 9-very probable) and “Given the information in the product webpage, the likelihood of purchasing the product is” (1-very low; 9-very high). *Perceived product quality* (α = 0.60) was measured using a single item: “What do you think is the quality of the isoTech tablet, overall?” (1-very low quality; 9-very high quality). The second product quality item was removed because the overall alpha was insufficient. An amended multi-item *perceived product quality* scale will be used in Study Two and Study Three.

As additional variables of interest, product category knowledge, brand familiarity, attitude toward the product webpage, use and trust of star ratings to evaluate a product, and the commonality and comfortability in using 5- or 10-point rating scale lengths to evaluate a product were measured. *Product category knowledge* ($\alpha = 0.76$) was measured using a 4-item scale: "I know pretty much about tablets", "Compared to most other people, I know less about tablets", "I am very knowledgeable about the product category of tablets" and "I do not feel very knowledgeable about tablets" (1-strongly disagree; 9-strongly agree). *Brand familiarity* was measured using a single item: "How familiar are you with the isoTech brand of tablets?" (1-not at all familiar; 9-very familiar). Following Chandran and Menon (2004), *overall product webpage attitude* ($\alpha = 0.95$) was assessed using a multi-scale item: "What is your overall attitude toward the product webpage?" (1-unfavorable; 9-favorable, 1-bad; 9-good, 1-negative; 9-positive). Use and trust of star ratings to evaluate products are measured using single items: "I typically review the consumer star ratings before making online purchases" and "When making online purchases, I trust the consumer star ratings to guide my decisions", respectively (1-strongly disagree; 9-strongly agree). *Scale length commonality* was measured using a single item: "The 5-star (10-star) scale in the ads is commonly used" (1-strongly disagree; 9-strongly agree). *Scale length comfort* was also measured using a single item: "I feel comfortable using the 5-star (10-star) rating scale to guide my online purchase decisions" (1-strongly disagree; 9-strongly agree). This item was used as the covariate because processing fluency literature suggests that higher levels of processing fluency contribute to more positive evaluations of a target (Winkielman et al 2003). Moreover, high fluency is more likely to exist when an individual is more familiar or more comfortable with the stimulus, like a consumer providing more favorable/positive product evaluations using a rating scale length that they are more comfortable with, for example (Alter and Oppenheimer, 2008). Per the criteria suggested by Nunnally (1978), scale reliability levels are satisfactory at Cronbach alpha levels at or greater than 0.70. For a complete listing of the items used, please refer to Appendix B for the full Study One stimulus.

Results

Manipulation and Other Checks

In each of the four conditions, participants responded to a manipulation check item: "According to the ad, what was the star rating of the product?" (participants responded using a text box) and an attention check item: "Select 8 on the scale below to demonstrate that you are reading carefully" (endpoints: strongly disagree-strongly agree). Frequency analysis revealed 41 out of 47 respondents (that is, 87%) in the 4 out of 5-point scale length condition qualified the manipulation and attention check; 43 out of 47 respondents (that is, 91%) in the 3.5 out of 5-point scale length condition qualified the manipulation and attention check; 39 out of 43 respondents (that is, 90%) in the 7 out of 10-point scale length condition qualified the manipulation and attention check; and 40 out of 43 respondents (that is, 93%) in the 8 out of 10-point scale length condition qualified the manipulation and attention check. Eliminating respondents who failed to qualify the manipulation and attention check items resulted in a sample of 163 respondents.

Since consumers are less likely to rely on the OCR ratings of recognized brands (Zhu and Zhang, 2010), the stimuli were designed using a generic brand of electronic tablet. Results of a one-sample t-test found that perception of *brand familiarity* for the electronic tablet was largely unfamiliar, compared to the scale median value of 5 ($M = 2.3$; $t = -23.71$; $p < 0.001$). General *product category knowledge* for electronic tablets was also significantly higher than the scale median value of 5 ($M = 5.83$, $p < .001$). In addition, respondents have a generally positive *attitude toward the webpage*, compared with the scale median value of 5 ($M = 5.96$; $t = 10.52$; $p < 0.001$).

Although both rating scale lengths can be seen in the digital marketplace, 5-point rating scales are more common. According to an independent samples t-test of a response to a single item question regarding commonality of the scale lengths, respondents confirm this substantive fact ($M_5 = 7.79$; $M_{10} = 3.55$; $t = 11.47$; $p < 0.001$). In general, participants report their use and trust of star rating scales in their online shopping experiences greater than the scale median value of 5 ($M_{StarUse} = 7.36$; $t = 15.72$; $p < 0.001$; $M_{StarTrust} = 4.83$; $t = 9.47$; $p < 0.001$). However, another independent samples t-test reveals that respondents feel more comfortable using the 5-point rating scale compared to the 10-point rating scale ($M_5 = 7.45$; $M_{10} = 4.83$; $t = 7.05$; $p < 0.001$). For this reason, *scale length comfort* was used as the covariate in the ANCOVA model to better investigate the comparative effects of the different rating scale lengths.

Hypothesis Test

ANOVA results revealed that the interaction effect between rating percentage and scale length was not significant for either *perceived product quality* ($F = 0.54$; $p = .46$) or *purchase intention* ($F = 0.39$; $p = .54$). There was a positive effect of rating percentage on *perceived product quality* ($F = 5.84$; $p = .017$) and on *purchase intention*, although it is marginally significant ($F = 3.19$; $p = .076$). Also, the effect of scale length on *perceived product quality* was marginally significant ($F = 3.28$; $p = .072$) and there was no effect of scale length on *purchase intention* ($F = 0.96$; $p = .328$).

A one-way analysis of covariance (ANCOVA) was conducted to determine a statistically significant difference between summary rating and scale length on *perceived product quality* and *purchase intention*, controlling for *scale length comfort*. Significant main effects for the ANCOVA results can be seen in Table 5.1.

Table 5.1: The Effect of Rating Percentage and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

Sources	***df***	***F-value (p-value)***	
		PQ	***PI***
Main Effects			
Rating Percentage	**1**	**5.09 (0.025)**	**2.65 (0.10)**
Scale Length	**1**	**7.22 (0.008)**	**3.56 (0.06)**
Scale Length Comfort	**1**	**5.22 (.024)**	**4.69 (.032)**
Interaction			
Rating Percentage*Scale Length	**1**	**0.44 (0.51)**	**0.31 (0.58)**
Residual	**159**		

Analysis based on the ANCOVA revealed that the interaction effects between rating percentage and scale length were not significant for either *perceived product quality* ($F = 0.44$; $p = .51$) or *purchase intention* ($F = 0.31$; $p = .58$). However, when positive ratings are displayed on a 10-point scale (i.e., 7 out of 10 and 8 out of 10) respondents' perception of product quality and purchase intentions are significantly greater than equally proportionate ratings on a 5-point scale (i.e., 3.5 out of 5 and 4 out of 5). The summary rating percentage had a significant main effect on *perceived product quality* ($M_{PQ70} = 4.84$; $M_{PQ80} = 5.41$; $F = 5.22$; $p = .024$) and a marginally significant main effect on *purchase intention* ($M_{PI70} = 3.92$; $M_{PI80} = 4.38$; $F = 2.64$ $p = .10$). Furthermore, results indicate that scale length had a significant main effect on *perceived product* *quality* ($M_{PQ5\text{-point scale}} = 4.74$; $M_{PQ10\text{-point scale}} = 5.51$; $F = 7.22$; $p = .008$) and a marginally significant main effect on *purchase intention* ($M_{PI5\text{-point scale}} = 3.84$; $M_{PI10\text{-point scale}} = 4.45$; $F = 3.56$; $p = .06$). The results were seen for both dependent measures while controlling for *scale length comfort* ($F_{Quality} = 5.22$; $p = 0.024$; $F_{PI} = 4.69$; $p = 0.032$).

Overall, higher summary ratings are positively related to *perceived product quality* and *purchase intentions,* in line with the findings from past research. Competing hypotheses predicted that a 10-point rating scale will either produce a lesser (H1) or greater (H1(ALT)) effect than a 5-point scale, depending on the scale endpoint that consumers use as an anchor. Results indicate that perceived product quality and purchase intention was higher when the star rating is on 10-point scale compared to 5-point, in support of hypothesis H1(ALT) and left-endpoint anchoring. Cell means for the effect of summary rating and scale length for both dependent variables can be viewed in Table 5.2. Additionally, to better visualize the differences in the cell mean summaries, means plots are provided for the effects on each dependent measure (see Figures 5.1 and 5.2).

Table 5.2: Cell Means for Effect of Rating Percentage and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

	70%		***80%***		***Overall Mean***	
Scale Length / Rating Percentage	***PQ***	***PI***	***PQ***	***PI***	***PQ***	***PI***
5-point	**4.55**	**3.69**	**4.95**	**4.00**	**4.74**	**3.84**
10-point	**5.14**	**4.14**	**5.88**	**4.76**	**5.51**	**4.45**
Overall Mean	**4.84**	**3.92**	**5.41**	**4.38**		

Figure 5.1: Effect of Rating Percentage and Scale Length on Perceived Product Quality

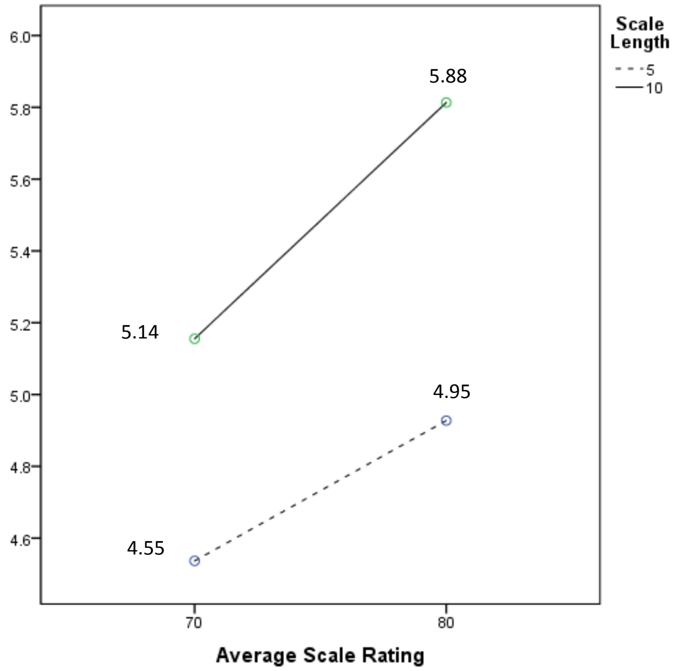
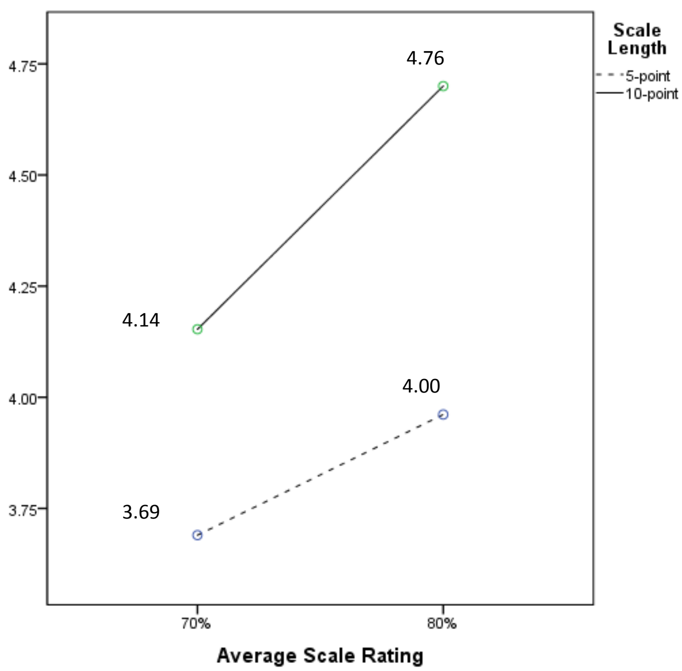


Figure 5.2: Effect of Rating Percentage and Scale Length on Purchase Intention



Study Two will explore the presence of a rating percentage display as a potential boundary condition to test when the effect of rating scale length on perceived product quality and purchase intention may diminish. A rating percentage displayed alongside the summary rating scale is a substantive quantitative factor that is currently seen in the OCR marketplace and may influence consumer perceptions of product quality and purchase intention.

CHAPTER 6: STUDY TWO

In Study One, we concluded that respondents seemingly anchor on the left-endpoint of the rating scale (H1(ALT)). Past research indicates the possibility that heuristic processing is a result of an individual's attempts to avoid numeral-based processing of the summary rating (Bonato et al. 2007). We test consumers' proneness to resorting to this heuristic further in Study Two by introducing a condition that is likely to obviate heuristic processing. The purpose of this study is to see how participants respond to ratings on 5- and 10-point scales, when the rating percentage is displayed next to the rating scale (e.g., ★★★★☆ 80%). A rating percentage is different from a summary rating, or valence, as discussed previously. Summary ratings are the numeric rating values (e.g., "4" out of 5 or "8" out of 10) whereas the rating percentage is computed by converting the summary rating to percentage (e.g., 4 out of 5 is 80%). Some online retailers display this percentage next to the rating scale (e.g., Rotten Tomatoes; Renters Voice) while others do not (e.g., Amazon; Zappos). Given our conceptual background discussion on heuristic processing, it seems beneficial to test the effects of displaying a rating percentage alongside the summary ratings on product evaluations because presence of rating percentage information might comprise a boundary condition to the effect observed in Study One.

Our theoretical premise is that people prefer concrete information and resort to heuristic processing only when adequate concrete information is absent. Since the percentage-based information is a more concrete representation of the rating score, respondents may be less likely to rely on the visual scale and hence resort to heuristic processing to evaluate the summary rating. In other words, we expect that the presentation of the rating in a more concrete form (e.g., as a percentage) removes the need to use visual processing heuristic, because the percentage-based information is likely to be interpreted uniformly by all, invariantly across scales of different lengths. Consequently, perceptual or behavioral outcomes are not likely to be a function of scale length when percentage rating information is present. This implies that perceived quality and purchase intention for a product would (not) be higher when a numerical summary rating is presented for a 10-point scale compared to when the same rating is presented on a 5-point scale, when rating percentage information is absent (present).

H2: When a rating percentage is absent, perceived product quality and purchase intention will be higher for a 10-point rating scale compared to a 5-point rating scale. When a rating percentage is present, scale length will have no effect on perceived product quality or purchase intention.

Methodology

Experimental Design

The experiment for Study Two involves a 2 (Scale length: 5-point; 10-point) x 2 (Rating percentage display: Present; Absent) between-subjects design. In this study, respondents in two of the four conditions will see a rating percentage numerically displayed next to the rating scale. This rating percentage provides the respondent with a correct computation of the visual star rating proportion. Like Study One, the volume of ratings will be held constant across conditions, at 1,394 customer ratings, and will be displayed alongside the star summary rating scale. In Study One, similar effects were found for both summary rating levels (70% and 80%) and so this study will include only summary ratings at the 80% level (i.e., 4 out of 5 and 8 out of 10), to simplify the design. The conditions related to the study will again be manipulated through an image of a product webpage like the pages displayed on Amazon, where potential buyers would typically view information about a product of interest. Please refer to Appendix C to view the product webpage stimuli used in this study.

Sample and Procedure

The procedure for this study is similar to Study One. One hundred and seventy-six undergraduate business students (Female = 47%; M_{Age} = 24) from a large Midwestern university participated in the experiment. After the participants were randomly assigned to one of four conditions, they first responded to some items that intended to measure their perceived product category knowledge. Next, each participant responded to questions related to perceived product quality and then purchase intentions. A modified multi-item measure of perceive product quality was introduced in this study, instead of the single-item measure used in Study One. In addition to the dependent variables of interest, we measured variables that customarily have been used in the OCR literature.

Measures

All variables were measured using 9-point Likert scales, unless otherwise noted. *Purchase intention* and *perceived product quality* are the dependent variables of interest. *Purchase intention* was measured using three items adapted from Keller et al. (1997), as in Study One (α = 0.91). *Perceived product quality* (α = 0.94) was measured using a multi-item scale adapted from Dodds et al. (1991): “The quality of this product seems to be” (1-very low; 9-very high), “This product seems to be reliable” (1-strongly disagree; 9-strongly agree), “The manufacturing quality of this product seems to be” (1-very low; 9-very high), “This product seems to be dependable” (1-strongly disagree; 9-strongly agree), and “This product is likely to be durable” (1-strongly disagree; 9-strongly agree).

Similar to Study One, we also measured *product category knowledge* (α = 0.82) and *attitude toward the webpage* (α = 0.96) using multi-item scales, and measured *brand familiarity* using a single item. Furthermore, we introduced *perceived product value*, *general attitude* *toward online ratings*, and a new multi-item scale on *scale length comfort* to replace the single-item measure in Study One. *Perceived product value* was measured using two items from Dodds et al. (1991): "The price shown for the product is" (1-very unacceptable; 9-very acceptable) and "This product is" (1-very poor value for the money; 9-very good value for the money) ($\alpha = 0.81$). *General attitude toward online ratings* ($\alpha = 0.85$) was measured using a multi-item scale adapted from Park et al. (2007) to replace the previous two items from Study One that asked respondents about their general use and trust of online ratings: "When I buy a product online, I always check the ratings that are presented on the website" (1-strongly disagree; 9-strongly agree), "When I buy a product online, the ratings presented on the website are helpful for my decision-making" (1-strongly disagree; 9-strongly agree), "When I buy a product online, the ratings presented on the website make me confident in purchasing the product" (1-strongly disagree; 9-strongly agree). *Scale length comfort* was measured using a multi-item scale adapted from Alter and Oppenheimer (2008) and included the scale commonality and scale comfortability items from Study One: "The 5-star (10-star) scale in the ads is commonly used" (1-strongly disagree; 9-strongly agree), "I feel comfortable using the 5-star (10-star) rating scale to guide my online purchase decisions" (1-strongly disagree; 9-strongly agree), "Please report the number of times you have seen a 5-star (10-star) rating scale used for online product ratings" (1-never seen before; 9-seen many times), and "I am familiar with the 5-star (10-star) rating scale used for online product ratings" (1-strongly disagree; 9-strongly agree) ($\alpha = 0.92$). Per the criteria suggested by Nunnally (1978), scale reliability levels are satisfactory at Cronbach alpha levels at or greater than 0.70, so all scales apparently meet this qualification. For a complete listing of the items used, please refer to Appendix D for the full study stimulus.

Results

Manipulation and Other Checks

In each of the four conditions, participants responded to a manipulation check item: "According to the product webpage, what was the star rating of the product?" (1-4 out of 5 stars; 2-8 out of 10 stars) and in the conditions where the rating percentage was displayed participants answered: "Was a rating percentage displayed next to the star rating in the product webpage?" (1-yes; 2-no) and if participants selected "yes" they answered "Was the displayed rating 80%?" (1-yes; 2-no). An attention check item was also presented: "Select 8 on the scale below to demonstrate that you are reading carefully" (endpoints: strongly disagree-strongly agree). Frequency analysis revealed 36 out of 44 respondents (that is, 82%) in the 5-point scale length with rating percentage "absent" condition qualified the manipulation and attention checks; 34 out of 44 respondents (that is, 77%) in the 5-point scale length with rating percentage "present" condition qualified the manipulation and attention checks; 35 out of 44 respondents (that is, 80%) in the 10-point scale length with rating percentage "absent" condition qualified the manipulation and attention check; and 36 out of 44 respondents (that is, 82%) in the 10-point scale length with rating percentage "present" condition qualified the manipulation and attention checks. Eliminating respondents who failed to qualify the manipulation and attention check items resulted in a sample of 141 respondents.

To eliminate the possible effects of using a recognized brand, we performed a one-sample t-test and found that the electronic tablet was largely perceived to be unfamiliar, given that the mean was below the scale median value of 5 ($M = 1.94$; $t = -22.56$; $p < 0.001$). General *product category knowledge* for electronic tablets was significantly higher than the scale median value of 5 ($M = 5.90$, $p < .001$), and so respondents seem to have some knowledge about the product they were evaluating. *Perceived value* for the tablet was found to be positive ($M = 6.48$; $t = 13.73$; $p < 0.001$) which means that respondents believed the advertised price was acceptable and the product was good value for the price. *General attitude toward online ratings* was similarly positive, compared with the scale median value of 5 ($M = 7.62$; $t = 22.52$; $p < 0.001$) indicating that the participants confidently use summary ratings before purchasing a product. In addition, respondents have a generally positive *attitude toward the webpage*, compared to scale median value of 5 ($M = 6.22$; $t = 9.06$; $p < 0.001$). According to an independent samples t-test of a response to the updated multi-item scale on comfortability regarding scale length (covariate), respondents again confirm that they are more comfortable using a 5-point rating scale compared to a 10-point rating scale ($M_{5\text{-point scale}} = 7.95$; $M_{10\text{-point scale}} = 3.08$; $t = 15.47$; $p < 0.001$).

Hypothesis Test

A one-way analysis of covariance (ANCOVA) was conducted to determine the effect of scale length and rating percentage on perceived product quality and purchase intention, controlling for the respondent's *scale length comfort*. The ANCOVA results are presented in Table 6.1.

Table 6.1: The Effect of Rating Percentage Display and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

Sources	***df***	***F-value (p-value)***	
		PQ	***PI***
Main Effects			
Scale Length	**1**	**9.44 (0.003)**	**6.88 (0.01)**
Rating Percentage Display	**1**	**0.92 (0.34)**	**0.82 (0.37)**
Scale Length Comfort	**1**	**14.30 (.001)**	**8.30 (.005)**
Interaction			
Rating Percentage Display*Scale Length	**1**	**1.16 (0.28)**	**1.58 (0.21)**
Residual	**137**		

Analysis based on the ANCOVA revealed that the interaction effects between scale length and rating percentage display were not significant for either *perceived product quality* ($F = 1.16$; $p = .28$) or *purchase intention* ($F = 1.58$; $p = .21$). However, as observed in Study One, we find that scale length had a significant main effect on *perceived product quality* and *purchase intention*. Thus, when ratings are displayed on a 10-point scale (i.e., 8 out of 10) respondents' perception of product quality and purchase intentions are higher than an equally proportionate rating on a 5-point scale (i.e., 4 out of 5): $M_{PQ5\text{-point scale}} = 5.87$; $M_{PQ10\text{-point scale}} = 6.90$; $F = 9.44$; $p = .003$; $M_{PI5\text{-point scale}} = 4.03$; $M_{PI10\text{-point scale}} = 5.40$; $F = 6.88$; $p = .01$, while controlling for the effect of a respondent's *scale length comfort* in evaluating products using either a 5-point or a 10-point rating scale length ($F_{Quality} = 14.30$; $p < 0.001$; $F_{PI} = 8.30$; $p = 0.005$). Significant main effects were not found for *rating percentage display* on *perceived product quality* ($M_{PQ\%Absent} =$

6.27; $M_{PQ\%Present}$= 6.50; F = 0.92; p = .34) or *purchase intention* ($M_{PI\%Absent}$= 4.55; $M_{PI\%P}$ = 4.87; F = 0.82; p = .37).

Cell means for the effect of scale length and rating percentage display for both dependent variables can be viewed in Table 6.2. Furthermore, to better visualize the differences in the cell mean summaries, means plots are provided for the effects on each dependent measure (see Figures 6.1 and 6.2).

Table 6.2: Cell Means for Effect of Rating Percentage Display and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

	%NP		***%P***		***Overall Mean***	
Scale Length / Rating Percentage Display	***PQ***	***PI***	***PQ***	***PI***	***PQ***	***PI***
5-point	**5.84**	**4.06**	**5.91**	**4.00**	**5.87**	**4.03**
10-point	**6.71**	**5.05**	**7.10**	**5.74**	**6.90**	**5.40**
Overall Mean	**6.27**	**4.55**	**6.50**	**4.87**		

Figure 6.1: Effect of Rating Percentage Display and Scale Length on Perceived Product Quality

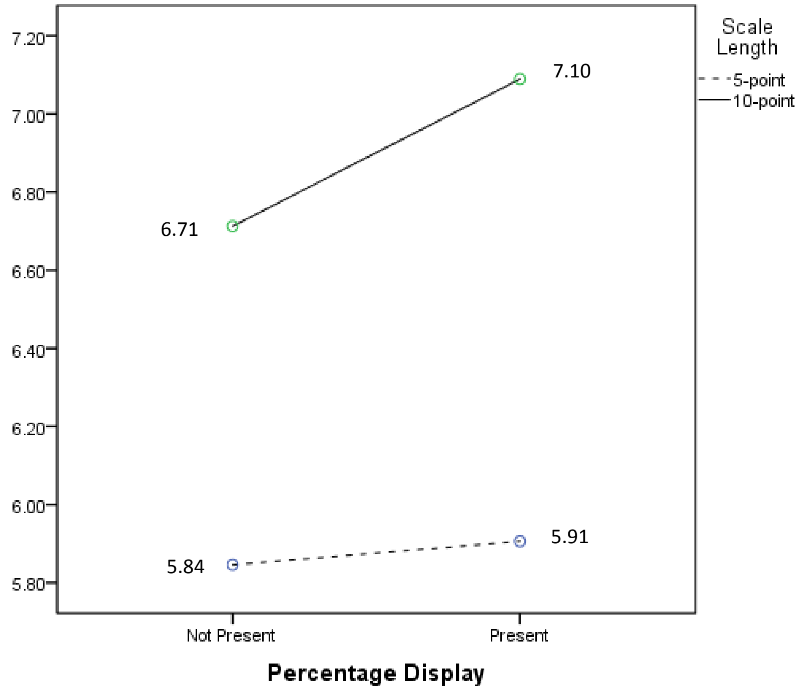
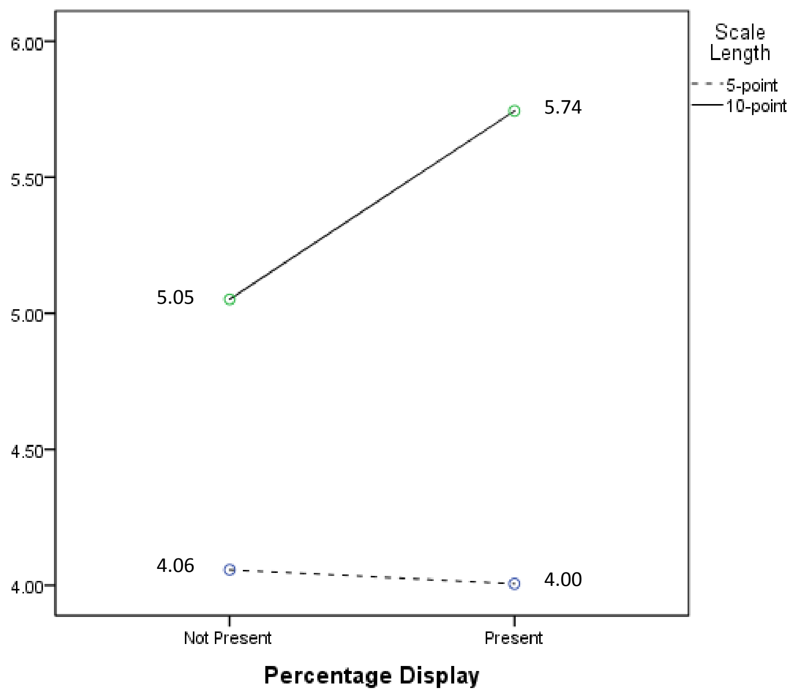


Figure 6.2: Effect of Rating Percentage Display and Scale Length on Purchase Intention



Although the results of the ANCOVA do not support our expected interaction effect between scale length and the rating percentage display (H2), we found support for a greater effect of a 10-point summary rating on perceived product quality and purchase intention, compared to 5-point summary rating of the same proportional value. All conditions in Studies One and Two held the volume of consumer ratings constant (1,394). In the market, rating volume varies drastically from a few ratings to tens of thousands of ratings. Study Three will explore rating volume as another potential boundary condition.

CHAPTER 7: STUDY THREE

The purpose of this study is to test whether volume of ratings, i.e., the number of ratings based on which the average rating is computed, acts as a boundary condition to the effects of scale length observed in the previous studies. Volume of ratings is the total number of individual customers whose ratings are aggregated to produce the summary rating. Research has shown a positive effect of volume of ratings on product evaluations. Thus, rating volume's positive effect can be seen but only when the summary rating is positive (Sun, 2012) or the brand of the product is weak or unfamiliar (Ho-Dac et al. 2013). These findings lead us to believe that our unfamiliar brand of tablet is an ideal setting in which to test the interaction effects of volume and scale length for positive summary ratings.

Thus far, the effects of scale length on perceived quality and purchase intention has been seen for high levels of rating volume (1,394) in the previous studies. This is in line with previous research, like Khare et al. (2011) who state that when an opinion is expressed by more and more people it is difficult to ignore. Just as an opinion expressed by many people might be difficult to ignore, that expressed by only a few might not be useful in formulating judgment. Thus, a low volume of ratings might not make the summary rating informative enough for consumers to base their perception of quality on the scale. As a decision-aid, the summary rating provided by the scale ought to be more reliable as the volume of ratings increases; so, if too few people have provided ratings then the rating scale may not be useful in judging the product. It is reasonable to posit that a certain threshold volume exists below which consumers will not rely on the scale as a decision-aid. In the context of our research, we label any volume of ratings that falls below this threshold as "low". We posit that at low levels of rating volume, a rating scale would be ignored as a decision-aid, regardless of its length. Consequently, perception of quality and purchase intention would not depend on scale length when rating volume is low. However, as rating volume increases beyond the low threshold value, the rating scale is used as a decision-aid, and consequently the effects of scale on perceived quality and purchase intention observed in the previous studies would occur. Formally:

> ***H3: When rating volume is high or medium, perceived product quality and purchase intention will be higher for a 10-point rating scale compared to a 5-point rating scale. However, the effect between 10- and 5-point rating scales will be absent when the rating volume is low.***

Methodology

Experimental Design

Study Three involves a 2 (Scale Length: 5-point; 10-point) x 3 (Rating Volume: Low; Medium; High) between-subjects experiment. Rating percentages were not displayed next to the scale and only summary ratings at the 80% level (i.e., 4 out of 5 and 8 out of 10) were included. The conditions related to the study were manipulated through a product webpage like the previous studies. Rating volume was a three-level factor, with low, medium, and high consumer rating volumes at 15, 90, and 1,394 individuals, respectively. Thirty undergraduate and graduate business students from a large Midwestern university were used as participants for a pretest to identify the three levels of rating volume. Respondents were first asked: “Imagine you were viewing a webpage to evaluate an electronic tablet before purchasing. For the rating score to be useful in your evaluation of the product, do you prefer there to be a minimum number of consumer ratings?” (1-yes; 2-no). Nearly ninety-seven percent (29 out of 30) of the sample answered the first question in the affirmative. If they selected “yes” to the first item they were directed to the second and final question: “Typically, what is the minimum number of consumer ratings that you think should be present, so that the rating score provided is useful to you in evaluating the tablet? [For instance, if you write a number "X" below, it means that the rating score provided is not useful if less than "X" people have rated the tablet.]" (text entry answer required). Sixty-five percent of the sample preferred a minimum rating volume of 15 individuals (35th percentile) before they would consider the rating score to be useful, and twenty-five percent of the sample preferred a minimum rating volume of 90 (75th percentile). The largest minimum rating volume in the sample was 200 ratings, so it is safe to assume that 1,394 can adequately represent the high rating volume condition. For robustness, the pretest questions were repeated for the same sample using two additional products (movie tickets and shoes) and the results supported the selected rating volume levels.

Sample and Procedure

Three hundred and fifty-five participants were recruited from Amazon Mechanical Turk (Female = 53%, M_{Age} = 35) to participate in Study Three and were randomly assigned to one of the six experimental conditions. All other procedures regarding the stimuli and measures were similar to the previous studies in that participants reviewed an image of a product webpage for an electronic tablet and then responded to all measures. Please refer to Appendix E for the product webpage stimuli used in this study.

Measures

All variables were measured using 9-point Likert scales, unless otherwise noted, like the previous studies. *Perceived product quality* ($\alpha = 0.95$) and *purchase intention* ($\alpha = 0.97$) were the dependent variables of interest. As in the previous studies, *perceived product quality* was measured using a multi-item scale from Dodds et al. (1991) and *purchase intention* was measured using three items adapted from Keller et al. (1997).

We measured *product category knowledge* (α = 0.78) and *attitude toward the webpage* (α = 0.95) using multi-item scales, and measured *brand familiarity* using a single item similar to Study One and Two. *Perceived product value* (α = 0.88) was measured using two items taken from Dodds et al. (1991) and *general attitude toward online ratings* (α = 0.88) was measured using a multi-item scale adapted from Park et al. (2007), as in Study Two. *Scale length comfort* (α = 0.94) was measured using a multi-item scale adapted from Alter and Oppenheimer (2008), similar to Study Two. Per the criteria suggested by Nunnally (1978), scale reliability levels were satisfactory at Cronbach alpha levels at or greater than 0.70. For a complete listing of the items used, please refer to Appendix F.

Results

Manipulation and Other Checks

In each of the six conditions, participants responded to a scale length manipulation check item: "According to the product webpage, what was the star rating of the product?" (4 out of 5 stars; 8 out of 10 stars) and a rating volume manipulation check item: "According to the product webpage, what was the number of consumer ratings for the product?" (15; 90; 1,395). An attention check item was also presented: "Select 8 on the scale below to demonstrate that you are reading carefully" (endpoints: strongly disagree-strongly agree). For the 5-point scale length conditions, frequency analysis revealed 53 out of 61 respondents (that is, 87%) in the low rating volume condition qualified the manipulation and attention check; 52 out of 58 respondents (that is, 90%) in the medium rating volume condition qualified the manipulation and attention check; and 50 out of 57 respondents (that is, 88%) in the high rating volume condition qualified the manipulation and attention check. For the 10-point scale length conditions, 46 out of 60 respondents (that is, 77%) in the low rating volume condition qualified the manipulation and attention check; 39 out of 57 respondents (that is, 68%) in the medium rating volume condition qualified the manipulation and attention check; and 44 out of 62 respondents (that is, 71%) in the high rating volume condition qualified the manipulation and attention check. Eliminating respondents who failed to qualify the manipulation and attention check items resulted in a sample of 284 respondents.

To eliminate the possible effects of using a recognized brand, we performed a one-sample t-test and found that the electronic tablet was largely perceived to be unfamiliar, given that the mean was below the scale median value of 5 ($M = 2.25$; $t = -24.89$; $p < 0.001$). *General product category knowledge* for electronic tablets was significantly higher than the scale median value of 5 ($M = 6.04$; $t = 12.11$; $p < .001$), and so respondents seem to have some knowledge about the product they were evaluating. *Perceived value* for the tablet was found to be positive ($M = 6.71$; $t = 18.48$; $p < 0.001$) which means that respondents believed the advertised price was acceptable and the product was good value for the price. *General attitude toward online ratings* was similarly positive, compared with the scale median value of 5 ($M = 7.54$; $t = 30.42$; $p < 0.001$), indicating that the respondents confidently use summary ratings before purchasing a product. In addition, respondents have a generally positive *attitude toward the webpage*, compared to a scale median value of 5 ($M = 6.82$; $t = 20.97$; $p < 0.001$). As in the previous studies, respondents confirmed that they were more comfortable using a 5-point rating scale compared to a 10-point rating scale ($M_{\text{5-point scale}} = 8.14$; $M_{\text{10-point scale}} = 4.42$; $t = 17.86$; $p < 0.001$).

Hypothesis Test

An analysis of covariance (ANCOVA) was conducted to determine the effect of scale length and volume of consumer ratings on *perceived product quality* and *purchase intention*, controlling for participants' *scale length comfort*. The ANCOVA results are presented in Table 7.1.

Table 7.1: The Effect of Rating Volume and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

Sources	***df***	***F-value (p-value)***	
		PQ	***PI***
Main Effects			
Scale Length	**1**	**9.64 (0.002)**	**4.01 (0.046)**
Rating Volume	**2**	**1.42 (0.243)***	**2.87 (0.058)***
Scale Length Comfort	**1**	**10.69 (.001)**	**1.76 (.186)***
Interaction			
Rating Volume*Scale Length	**2**	**.34 (0.711)**	**.21 (0.812)**
Residual	**278**		

* These statistics were reported using the multi-item scale length comfort measure. However, ANCOVA results improve when the original single-item scale length comfort measure is used: effect of rating volume on product perceived quality is still not statistically significant but does improve ($F = 1.91$; $p = 0.150$); the effect of rating volume on purchase intention becomes statistically significant ($F = 3.03$; $p = 0.50$); and the effect of scale length comfort on purchase intention becomes statistically significant ($F = 4.62$; $p = 0.032$).

Contrary to H3, analysis based on the ANCOVA revealed that the interaction effect between scale length and volume of consumer ratings was not significant for either perceived product quality ($F = .34$; $p = .71$) or purchase intention ($F = .21$; $p = .81$). However, we again find a significant effect of scale length. When ratings are displayed on a 10-point scale (i.e., 8 out of 10) respondents' perception of product quality and purchase intentions are significantly greater than an equally proportionate rating on a 5-point scale (i.e., 4 out of 5). Our results indicate that scale length had a significant main effect on perceived product quality ($M_{PQ5\text{-point}}$

$_{scale}$ = 6.25; $M_{PQ10\text{-point scale}}$ = 6.92; F = 9.64; p = .002) and purchase intention ($M_{PI5\text{-point scale}}$ = 4.95; $M_{PI10\text{-point scale}}$ = 5.65; F = 4.01; p = .046), while controlling for the effect of a respondent's *scale length comfort* in evaluating products using either a 5-point or a 10-point rating scale length ($F_{Quality}$ = 10.67; p = 0.001; F_{PI} = 1.76; p = 0.186). Significant main effects were not found for volume of consumer ratings on perceived product quality ($M_{PQLow\ volume}$ = 6.43; $M_{PQMed\ volume}$ = 6.58; $M_{PQHigh\ volume}$ = 6.74; F = 1.42; p = .243) but were found for purchase intention at a marginal level ($M_{PQLow\ volume}$ = 4.88; $M_{PQMed\ volume}$ = 5.44; $M_{PQHigh\ volume}$ = 5.58; F = 2.87 p = .058). Contrast effects among volume of consumer rating levels for purchase intention reveal a statistically significant difference between low and high consumer rating levels (p = .024), but not between low and medium consumer rating levels (p = .077) or between medium and high consumer rating levels (p = .644).

Cell means for the effect of scale length and volume of consumer ratings for both dependent variables can be viewed in Table 7.2. To showcase the differences in the cell mean summaries, means plots are again provided for the effects on each dependent measure (see Figures 7.1 and 7.2).

Table 7.2: Cell Means for Effect of Rating Volume and Scale Length on Perceived Product Quality (PQ) and Purchase Intention (PI)

	Low		***Medium***		***High***		***Overall Mean***	
Scale Length / Rating Volume	***PQ***	***PI***	***PQ***	***PI***	***PQ***	***PI***	***PQ***	***PI***
5-point	**6.14**	**4.62**	**6.16**	**4.98**	**6.46**	**5.24**	**6.25**	**4.95**
10-point	**6.72**	**5.14**	**7.01**	**5.90**	**7.03**	**5.92**	**6.92**	**5.65**
Overall Mean	**6.43**	**4.88**	**6.58**	**5.44**	**6.74**	**5.58**		

Figure 7.1: Effect of Rating Volume and Scale Length on Perceived Product Quality

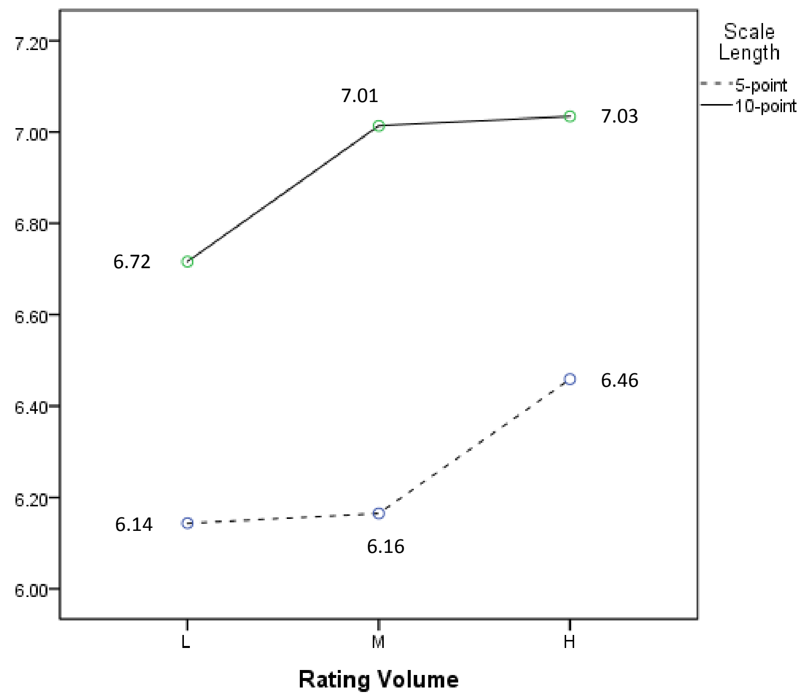
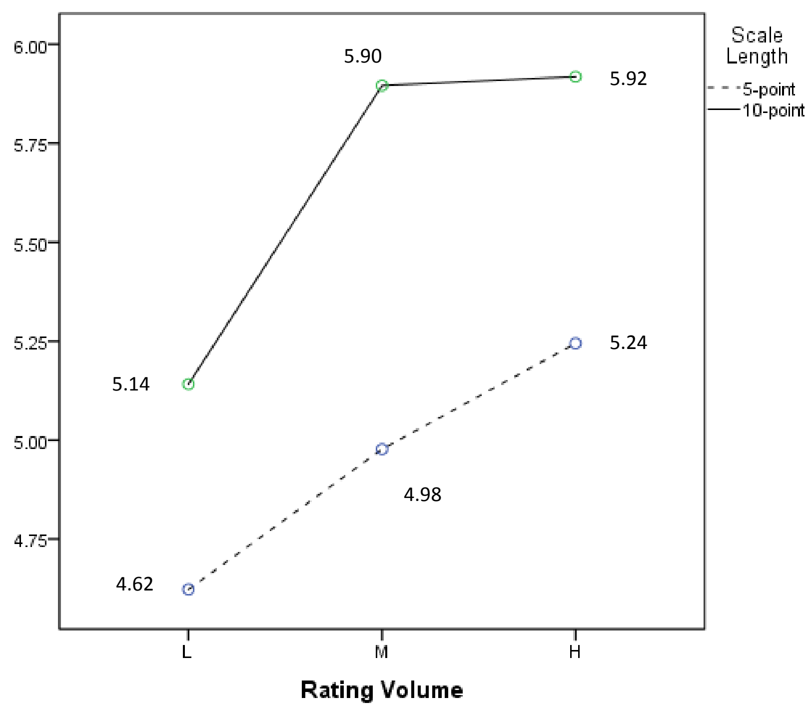


Figure 7.2: Effect of Rating Volume and Scale Length on Purchase Intention



The interaction between scale length and volume of consumer ratings was not significant, contrary to hypothesis H3. Additionally, there was no main effect of rating volume on perceived product quality. However, there was a positive main effect of rating volume on purchase intention, specifically the difference between low and high levels of rating volume. However, Study Three results indicate that the effect of a 10-point summary rating on *perceived product quality* and *purchase intention* is greater than a 5-point summary rating for the same rating proportion.

CHAPTER 8: DISCUSSION

The focus of this research was to test the effects of OCR summary rating scale lengths on perceived product quality and purchase intention. Quantitative OCRs are scale ratings from individual consumers that are aggregated into a summary statistic to be displayed on a product webpage. The OCR literature generally finds that summary ratings and consumer rating volume have positive effects on product evaluations. Although researchers have tested OCR effects using 5- and 10-point rating scales separately, they have not investigated the effect of rating scale length on product evaluations. According to scale length research, the optimal number of response categories could vary anywhere from 2 to 25, depending on individual preferences (Viswanathan et al. 2004). It seems reasonable to expect that if scale length can affect an individuals' provision of ratings, it might also influence their evaluations of others' ratings.

OCR summary ratings are proportions (e.g., 4 out of 5) and Bonato et al. (2007) state that it may be difficult for individuals to think in terms of proportions. When individuals have real or perceived difficulty interpreting the summary rating in terms of a proportion involving numerals (numeral-based processing) of the summary rating, they rely on heuristics, like anchoring, to reach a conclusion (Bagchi and Davis, 2012). Furthermore, Sherif and Hovland (1961) claim that the end values of a range usually acquire an anchoring role, so individuals may either anchor on the left or right-endpoint of a summary rating scale. For this reason, competing hypotheses about the left and right-endpoints of the summary rating scales were presented. Three studies were performed to test the effects of summary rating scale length on perceived product quality and purchase intention, with potential moderators.

Study One tested competing hypotheses to explore the potential differences in consumers' evaluation of a product using a summary rating scale of different lengths (5-point and 10-point). Results support that perceived product quality and purchase intention will be higher when the summary rating is on a 10-point scale compared to a 5-point scale. Since the 5- and 10-point summary ratings provided to respondents were proportionately equivalent (70% and 80% were the two ratings used in this study), we inferred from this finding that respondents seem to be evaluating the scale via a visual processing heuristic, viz., anchoring on the left-endpoint of the scale. Anchoring on the left-endpoint produces more favorable (weaker) product evaluations for the 10-point (5-point) rating scale as the distance from the left-endpoint to the summary rating score is greater (weaker), despite equal proportions between the ratings on the different scales, both quantitatively and geometrically. Thus, a rating of 8 on a 10-point scale is visually farther away from the left-endpoint “1” than a rating of 4 on a 5-point scale is from its left-endpoint “1”, holding the sizes of the stars constant across the two scales. Instead of interpreting the summary rating proportions numerically and arriving at similar values (e.g., 4/5 = 80% and 8/10 = 80%), consumers appear to be using use visual heuristics which facilitate evaluation potentially at the cost of evaluative accuracy. Thus, although from a rational standpoint, perceptual outcomes should not differ across the scale lengths for the same rating, visual processing heuristic based on left-endpoint anchoring leads to differences.

In Study Two, a boundary condition was introduced to potentially remove the need for heuristic processing. The purpose of this study was to test if differences in perceived product quality and purchase intention across the two different scale lengths observed in Study 1 would disappear when a rating percentage was displayed next to the rating scales. Since the rating percentage presents the summary rating in a form that is invariant across the two scales, i.e., 80% for both a 4(8) out 5 (10) rating, it was hypothesized that respondents would be less likely to rely on heuristics to evaluate the summary rating when the rating percentage was present (i.e., scale length would have no effect on perceptual outcomes). However, results indicated reliance on left-endpoint-based visual-processing heuristic despite the presence of percentage rating. Thus, perceived quality and purchase intention were higher for the 10-point scale than for the 5-point scale, regardless of the provision of percentage rating. Despite the potential for respondents to not rely on the summary rating in the presence of more concrete information like a rating percentage display, left anchoring via visual heuristic processing is not abandoned but remains an overriding force which influences perceptions of product quality and purchase intention. This finding speaks to the inherently powerful influence of this processing heuristic on consumers' use of rating scales.

The purpose of Study Three was to test the effects of rating scale length on perceived product quality and purchase intention, across various levels of rating volume. It was hypothesized that at high (1,394) and medium (90) volume levels the effect of scale length on perceived quality and purchase intention would be greater for the 10-point summary rating, compared to the 5-point summary rating, but at low levels of rating volume (15) the effect of scale length would be absent. Although the predicted interaction between scale length and volume of consumer ratings was not significant, results indicate that across all levels of rating volume the effect of a 10-point summary rating on perceived product quality and purchase intention is greater than a 5-point summary rating of the same proportion. There was also a main effect of volume on purchase intention, indicating that as volume increases in level (from low to high) purchase intentions increase. Due to the low consensus of consumer evaluations in the low volume condition, summary ratings could arguably be easier to ignore as a product evaluation tool and decision aid. However, respondents appear to be evaluating the summary ratings heuristically, given the greater effect of the 10-point scale, compared to the 5-point scale, even at low levels of consumer rating volume. We might recall that the low rating volume we used in Study Three was one that was reported by a large majority of the respondents in a preceding pretest to be a level at which rating scales were not useful to them for evaluating products. Thus, at this rating volume, most respondents would admittedly ignore rating scales. However, our findings show that they are unable to escape the influence of the powerful visual-processing heuristic based on left-endpoint anchoring, which presumably operates non-consciously. All in all, Studies Two and Three indicate that visual-processing heuristic based on the left-endpoint anchor is a powerful non-conscious heuristic that overrides conditions under which such heuristic should not be operational, if processing of scale-based summary ratings occurred entirely consciously and rationally.

In general, respondents appear to continue to rely on a visual cue (summary rating scale) even when presented with more concrete (rating percentage display) or more informative (rating volume) information. The question remains: *why do respondents evaluate the product and its summary rating using visual heuristics when additional information is presented?* One explanation is that words and numbers are processed sequentially, whereas an image can be processed more quickly and automatically (in gestalt fashion); furthermore, the connection between an image and its meaning is more direct and automatic than it is for words (Luna and Peracchio, 2003). Further support for the automatic and even unconscious nature of image processing can be found by observing how aesthetics influence perceptions of attractiveness (Townsend and Kahn, 2013). We originally hypothesized that visual heuristics would lead to scale endpoint anchoring as a shortcut to numeral-based processing. However, it is possible that the mere presence of a visual cue (rating scale image) trumps a rating percentage (numbers) and rating volume (numbers and words) because of the ease with which the visual image can be processed in comparison to numbers or words.

Practical implications of this research are that product summary ratings are a tool that consumers can use to communicate the quality of a product, potentially alleviating purchasing uncertainties for potential buyers. Star-rating platforms (e.g., Amazon) and crowd-source review sites (e.g., Yelp) are pervasive and play an increasingly significant role in today's online marketplace. Seemingly, retailers follow the lead of retailing giants, like Amazon, in how they organize and display their rating scales online. To produce more favorable product evaluations in a pre-purchase setting, retailers can display ratings of different scale lengths.

Limitations

This section will discuss potential limitations that generally apply across all three studies: the need for additional product testing and field studies, and the functionality of the product information webpage (stimuli). Additional limitations are identified that directly relate to the individual studies: the lack of negative summary ratings (Study One) and the need to test additional lower levels of rating volume (Study Three).

By introducing additional products in the research, not only may the validity of the results improve but possible effects of self-assessed product category knowledge and brand familiarity could be tested. Past research suggests that as an individual's self-assessed knowledge of a product category increases (decreases), they are less (more) likely to rely on available information in their evaluation of the product (Park et al. 1988). Because respondents in our studies reported higher levels of product category knowledge they may have ignored additional information (i.e., the rating percentage or rating volume) and instead processed the summary ratings heuristically. Additionally, in a real purchase situation, where stakes are higher for an actual consumer, individuals may pay more attention to concrete info and rely less on visual cues. A field study can provide valuable insights that are needed to validate the findings of these survey-based experiments in this research.

Another potential limitation is that the product information webpage is limited in its functionality for the respondents. Unlike an actual product webpage where consumers can interact with the OCR data and click to receive richer information on the rating distribution or read text-based reviews, the stimuli in all three studies was a static image. OCR research finds that most consumers prefer to read text-based reviews prior to purchase. Including text-based review analysis in future studies is complicated, however, it may be worthwhile to simply measure a respondent's preferences for viewing a rating distribution or reading text-based reviews.

In Study One, the effect of scale length on product evaluations is examined for positive summary ratings only. Examining the effects of scale length on product evaluations for neutral (e.g., 3 out of 5 or 6 out of 10) or negative summary ratings (e.g., 2 out of 5 or 4 out of 10) may provide additional understanding to the effect of scale length. Differences might appear from the increase in purchase risk, given a negative summary rating, or they may arise from the visual change in summary rating scale proportion. The proclivity to anchor on the left-endpoint of the scale may change when the summary ratings are less than positive, possibly altering the scale effect that has been observed in the present research. Study One included competing hypotheses, predicting greater effects for either scale length depending on the endpoint in which respondents' anchor. In the presence of negative summary ratings, respondent's anchor of choice may switch from the left-endpoint to the right-endpoint, in which case a 5-point scale rating would have a greater effect than a 10-point scale reporting the same summary rating.

Finally, for Study Three it was predicted that at lower levels of rating volume, respondents are less likely to rely on the summary rating as a decision aid, given the increased risk of relying on a quality signal that is a subjective opinion expressed by so few individuals. However, there is still a need to explore lower volume limits where a lack of scale effect might present itself. Although there does not appear to be an upper limit on consumer rating volume (i.e., effect of scale length was seen for 15, 90, and 1,394 consumer ratings), we have yet to find the lower limit, if one exists. Additionally, all three studies included the scale length comfort covariate in the analysis. The inconsistent results of Study Three may have occurred due to the instability of the covariate scale itself. For now, the single-item covariate scale seems to have served its purpose well. Moving forward, we plan to use a more rigorously tested multi-item scale.

Future Research

Proposed studies will serve as logical extensions of the current research and help to resolve the limitations that were identified. Additional research can again examine the effects of summary rating scale length on perceived product quality and purchase intention with a specific focus on (1) further testing of the visual heuristic based on left-endpoint anchoring; (2) additional product testing and field tests; (3) negative summary ratings, and; (4) additional lower levels of rating volume.

A secondary study could replace the two full-length star scales and replace them with rectangular bar (approximately the length of the original 5- and 10-point star rating scales, respectively) that are proportionately filled-in to represent a summary rating. By using a bar instead of stand-alone stars, respondents are unable to count or see individual stars and must evaluate the product with only the filled-in space representing a summary ratings and not numerical quantity on a scale.

Introducing additional products in future studies could add more validity to the findings. Furthermore, if products are selected from categories in which respondents have lower levels of perceived product category knowledge they may be more inclined to evaluate the summary rating more objectively and less heuristically. Thus, boundary conditions regarding product category knowledge may be identified where the 5- or 10-point scales do not differ in their effect on product evaluations. In addition, partnering with an organization who utilizes a star summary rating scale platform for their products to conduct a field study is necessary to test the effects of scale length in an actual consumption setting.

Sun (2012) indicated that nearly 65% of Amazon products were rated at a level of 4.1 out of 5 or higher. For this reason, it is reasonable to give attention to positive ratings only, as was done in the current research. However, it is important to test the effect of summary rating scale length on product evaluations for negative summary ratings as well. Thus, additional studies could examine negative summary rating values for both scale lengths, to see if the method in which consumers evaluate such values differs from the left-endpoint anchoring of the positively rated products seen in this research.

This research examined three levels of consumer rating volume (15, 90, and 1,394) across both scale lengths. However, the predicted interaction effect and predicted lack of scale length effect at low levels of volume were not observed. Effort could be directed toward further examination of the role of rating volume levels to explore the lower limits, specifically, below 15. Testing various levels to find a lower limit could be beneficial for newer products on the market that lack an established base of ratings or for less popular products that are experiencing slow sales. Furthermore, this research explored the moderating effect of rating volume for positive ratings only, but future studies should also test the effect of volume for negative ratings.

APPENDIX A: Product Webpage Stimuli (Study One)

3.5 out of 5 (70%)

4 out of 5 (80%)

7 out of 10 (70%)

8 out of 10 (80%)

APPENDIX B: Complete Stimulus for All Experimental Conditions (Study One)

Research Information Sheet

Title of Study: Star Rating Evaluation Survey

Principal Investigator (PI): Aaron Johnson (Marketing Department) 313-577-4406

Purpose: You are being asked to be in a research study of online products presented in a product webpage because you are a potential consumer of this type of product. Study Procedures: If you take part in the study, you will be asked to:

* Fill out a survey.

* Answer questions about your attitude toward the ad that you will view and also your online shopping experiences. In addition, you will be asked background questions on your demographic information. You may skip background related questions and still complete the survey.

* This is a one-time survey that should take approximately 10 minutes to complete.

Benefits: As a participant in this research study, there will be no direct benefit for you; however, information from this study may benefit other people now or in the future.

Risks: There are no known risks at this time to participation in this study.

Costs: There will be no costs to you for participation in this research study.

Compensation: For taking part in this research study, you will be paid for your time via bonus points awarded by your instructor.

Confidentiality: All information collected about you during the course of this study will be kept without any identifiers.

Voluntary Participation /Withdrawal: Taking part in this study is voluntary. You are free to withdraw at any time. Your decision will not change any present or future relationships with Wayne State University or its affiliates.

Questions: If you have any questions about this study now or in the future, you may contact Aaron Johnson at the following phone number 313-577-4406. If you have questions or concerns about your rights as a research participant, the Chair of the Institutional Review Board can be contacted at (313) 577-1628. If you are unable to contact the research staff, or if you want to talk to someone other than the research staff, you may also call the Wayne State Research Subject Advocate at (313) 577-1628 to discuss problems, obtain information, or offer input.

Participation: By completing the questionnaire you are agreeing to participate in this study. Click the "Next" button below to continue.

This survey will first have you respond to statements about an electronic tablet. Next, you will view a product webpage for the electronic tablet and answer the subsequent questions. There are no right or wrong answers, but we ask that you please take your time and answer truthfully. If you agree to participate, please click the button below to begin the survey.

Please indicate how much you agree or disagree with the following statements regarding electronic tablets:

I know pretty much about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Compared to most other people, I know less about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I am very knowledgeable about the product category of tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I do not feel very knowledgeable about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Please carefully review the product webpage below and then answer the questions on the following pages.

According to the product webpage, what was the star rating of the product?
(______________________)

What do you think is the quality of the isoTech® tablet, overall?
(Very Low Quality; 9-Very High Quality)

How attractive is the star rating for the advertised product?
(1-Not at All Attractive; 9-Very Attractive)

If you were going to buy a tablet, how likely would you be to buy the tablet shown in the product webpage?
(1-Not Likely at All; 2-Very Likely)

Given the information in the product webpage, the likelihood of purchasing this product is
(1-Very Low; 9-Very High)

How probable is it that you would consider purchasing this product?
(1-Not Probable; 2-Very Probable)

The next few questions ask about your responses to the product webpage you saw at the start of the survey.

How familiar are you with the isoTech® brand of tablets?
(1-Not at All Familiar; 9-Very Familiar)

How believable is the product webpage?
(1-Not at All Believable; 9-Highly Believable)
(1-Not at All Acceptable; 9-Highly Acceptable)

Do you feel that all the ad information (text and graphics) is congruent with your expectations?
(1-Totally Unexpected; 9-Totally Expected)
(1-Very Different; 9-Not at All Different)

Select 8 on the scale below to demonstrate that you are reading carefully.
(1-Strongly Disagree; 9-Strongly Agree)

How credible do you feel the product webpage is?
(1-Not at All Credible; 9-Very Credible)

How involved were you in analyzing the product webpage?
(1-Very Uninvolved; 9-Very Involved)
(1-Concentrated Very Little; 9-Concentrated Very Hard)
(1-Paid Very Little Attention; 9-Paid a lot of Attention)

The text in the product webpage was...
(1-Difficult to Process; 9-Easy to Process)
(1-Difficult to Understand; 9-Easy to Understand)

What is your overall attitude toward the product webpage?
(1-Unfavorable; 9-Favorable)
(1-Bad; 9-Good)
(1-Negative; 9-Positive)

Briefly share your thoughts about the star rating for the product and if/how it influenced your attitude toward the product.
(____)

The 5-star rating scale in the product webpage is commonly used. <u>OR</u> The 10-star rating scale in the ads is commonly used.
(1-Strongly disagree; 9-Strongly Agree)

I feel comfortable using the 5-star rating scale to guide my online purchase decisions. <u>OR</u> I feel comfortable using the 10-star rating scale to guide my online purchase decisions.
(1-Strongly disagree; 9-Strongly Agree)

I enjoy work that requires the use of numbers.
(1-Strongly disagree; 9-Strongly Agree)

I find it satisfying to solve day-to-day problems involving numbers.
(1-Strongly disagree; 9-Strongly Agree)

Numerical information is very useful in everyday life.
(1-Strongly disagree; 9-Strongly Agree)

I prefer not to pay attention to information involving numbers.
(1-Strongly disagree; 9-Strongly Agree)

I do not like to think about issues involving numbers.
(1-Strongly disagree; 9-Strongly Agree)

I like to make calculations using numerical information.
(1-Strongly disagree; 9-Strongly Agree)

I don't find numerical information to be relevant for most situations.
(1-Strongly disagree; 9-Strongly Agree)

I like to go over numbers in my mind.
(1-Strongly disagree; 9-Strongly Agree)

I typically review the consumer star ratings, before making online purchases.
(1-Strongly disagree; 9-Strongly Agree)

When making online purchases, I trust the consumer star ratings to guide my decisions.
(1-Strongly disagree; 9-Strongly Agree)

I review the consumer written reviews, before making online purchases.
(1-Strongly disagree; 9-Strongly Agree)

When making online purchases, I trust the consumer written reviews to guide my decisions.
(1-Strongly disagree; 9-Strongly Agree)

I am an experienced online shopper.
(1-Strongly disagree; 9-Strongly Agree)

On average, how often do you make online purchases each month?
(______________________________)

In your estimation, how many total times have you made an online purchase?
(______________________________)

What is your gender?
(1-Male; 2-Female)

What is your age?
(_______________)

What is your ethnicity?
(1-White; 2-Black or African American; 3-American Indian or Alaska Native; 4-Asian; 5-Native Hawaiian or Pacific Islander; 6-Other)

Thank you for participating in this survey. Click the "next" button to be redirected to a separate page where you can anonymously provide your name, and instructor's name, to receive bonus points for participation.

APPENDIX C: Product Webpage Stimuli (Study Two)

4 out of 5 (% Display Absent)

4 out of 5 (% Display Present)

8 out of 10 (% Display Absent)

All-new
isoTech Tab Pro HD8
12-hour battery | 2X the storage

isoTech Tab Pro Tablet, 8'' HD Display, 16GB, Wifi, Bluetooth 4.0 (Black)
by LabTech

★★★★★★★★☆☆ 1,394 customer reviews

Price: $89.99 & FREE Shipping.

Color: Black

- 1.6 GHz Quad-Core Processor
- Supports Wi-Fi connection
- Runs Android 5.1 (Lollipop) OS
- 1.5 GB RAM memory
- 8-inch HD display
- Screen resolution of 1280 x 800 pixels
- 16 GB of internal storage
- Micro SD slot for up to 200 GB of expandable storage
- Up to 12 hours of battery life
- 2 MP front camera and 5 MP rear camera

8 out of 10 (% Display Present)

All-new
isoTech Tab Pro HD8
12-hour battery | 2X the storage

isoTech Tab Pro Tablet, 8'' HD Display, 16GB, Wifi, Bluetooth 4.0 (Black)
by LabTech

★★★★★★★★☆☆ (80%) 1,394 customer reviews

Price: $89.99 & FREE Shipping.

Color: Black

- 1.6 GHz Quad-Core Processor
- Supports Wi-Fi connection
- Runs Android 5.1 (Lollipop) OS
- 1.5 GB RAM memory
- 8-inch HD display
- Screen resolution of 1280 x 800 pixels
- 16 GB of internal storage
- Micro SD slot for up to 200 GB of expandable storage
- Up to 12 hours of battery life
- 2 MP front camera and 5 MP rear camera

APPENDIX D: Complete Stimulus for All Experimental Conditions (Study Two)

Research Information Sheet

Title of Study: Star Rating Evaluation Survey

Principal Investigator (PI): Aaron Johnson (Marketing Department) 313-577-4406

Purpose: You are being asked to be in a research study of online products presented in an product webpage because you are a potential consumer of this type of product. Study Procedures: If you take part in the study, you will be asked to:

* Fill out a survey.

* Answer questions about your attitude toward the ad that you will view and also your online shopping experiences. In addition, you will be asked background questions on your demographic information. You may skip background related questions and still complete the survey.

* This is a one-time survey that should take approximately 10 minutes to complete.

Benefits: As a participant in this research study, there will be no direct benefit for you; however, information from this study may benefit other people now or in the future.

Risks: There are no known risks at this time to participation in this study.

Costs: There will be no costs to you for participation in this research study.

Compensation: For taking part in this research study, you will be paid for your time via bonus points awarded by your instructor.

Confidentiality: All information collected about you during the course of this study will be kept without any identifiers.

Voluntary Participation /Withdrawal: Taking part in this study is voluntary. You are free to withdraw at any time. Your decision will not change any present or future relationships with Wayne State University or its affiliates.

Questions: If you have any questions about this study now or in the future, you may contact Aaron Johnson at the following phone number 313-577-4406. If you have questions or concerns about your rights as a research participant, the Chair of the Institutional Review Board can be contacted at (313) 577-1628. If you are unable to contact the research staff, or if you want to talk to someone other than the research staff, you may also call the Wayne State Research Subject Advocate at (313) 577-1628 to discuss problems, obtain information, or offer input.

Participation: By completing the questionnaire you are agreeing to participate in this study. Click the "Next" button below to continue.

This survey will first have you respond to statements about an electronic tablet. Next, you will view an ad for the electronic tablet and answer the subsequent questions after each ad. There are no right or wrong answers, but we ask that you please take your time and answer truthfully. If you agree to participate, please click the button below to begin the survey.

Please indicate how much you agree or disagree with the following statements regarding electronic tablets:

I know pretty much about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Compared to most other people, I know less about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I am very knowledgeable about the product category of tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I do not feel very knowledgeable about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Please carefully review the product webpage below and then answer the questions on the following pages.

According to the product webpage, what was the star rating of the product?
(1-4 out of 5 stars; 2-8 out of 10 stars)

Was a rating percentage displayed next to the star rating in the product webpage?
(1-Yes; 2-No)

Was the displayed rating percentage 80%?
(1-Yes); 2-No)
The quality of this product seems to be
(1-Very Low; 9-Very High)

This product seems to be reliable.
(1-Strongly Disagree; 9-Strongly Agree)

The manufacturing quality of this product seems to be
(1-Very Low; 9-Very High)

This product seems to be dependable.
(1-Strongly Disagree; 9-Strongly Agree)

This product is likely to be durable.
(1-Strongly Disagree; 9-Strongly Agree)

If you were going to buy a tablet, how likely would you be to buy the tablet shown in the product webpage?
(1-Not Likely at All; 2-Very Likely)

Given the information in the product webpage, the likelihood of purchasing this product is
(1-Very Low; 9-Very High)

How probable is it that you would consider purchasing this product?
(1-Not Probable; 2-Very Probable)

The next few questions ask about your responses to the product webpage you saw at the start of the survey.

How familiar are you with the isoTech® brand of tablets?
(1-Not at All Familiar; 9-Very Familiar)

How believable is the product webpage?
(1-Not at All Believable; 9-Highly Believable)
(1-Not at All Acceptable; 9-Highly Acceptable)

What is your overall attitude toward the product webpage?
(1-Unfavorable; 9-Favorable)
(1-Bad; 9-Good)
(1-Negative; 9-Positive)

Select 8 on the scale below to demonstrate that you are reading carefully.
(1-Strongly Disagree; 9-Strongly Agree)

The price shown for the product is
(1-Very Unacceptable; 2-Very Acceptable)

This product is a
(1-Very Poor Value for the Money; 2-Very Good Value for the Money)

The 5-star rating scale in the product webpage is commonly used. <u>OR</u> The 10-star rating scale in the product webpage is commonly used.
(1-Strongly disagree; 9-Strongly Agree)

Please report the number of times you have seen a 5-star rating scale used for online product ratings. OR Please report the number of times you have seen a 10-star rating scale used for online product ratings.
(1-Never Seen Before; 2-Seen Many Times)

I am familiar with the 5-star rating scale used for online product ratings. OR I am familiar with the 10-star rating scale used for online product ratings.
(1-Strongly disagree; 9-Strongly Agree)

I feel comfortable using the 5-star rating scale to guide my online purchase decisions. OR I feel comfortable using the 10-star rating scale to guide my online purchase decisions.
(1-Strongly disagree; 9-Strongly Agree)

When I buy a product online, I always check the ratings that are presented on the website.
(1-Strongly disagree; 9-Strongly Agree)

When I buy a product online, the ratings presented on the website are helpful for my decision-making.
(1-Strongly disagree; 9-Strongly Agree)

When making online purchases, I trust the consumer star ratings to guide my decisions.
(1-Strongly disagree; 9-Strongly Agree)

When I buy a product online, the ratings presented on the website make me confident in purchasing the product.
(1-Strongly disagree; 9-Strongly Agree)

I am an experienced online shopper.
(1-Strongly disagree; 9-Strongly Agree)

On average, how often do you make online purchases each month?
(______________________________)

Do you currently own an electronic tablet?
(1-Yes; 2-No)

Are you currently looking to purchase an electronic tablet?
(1-Yes; 2-No)

What operating system platform would you prefer to use in an electronic tablet?
(1-Apple iOS; 2-Android; 3-Other)

What is your gender?
(1-Male; 2-Female)

What is your age?
(______________)

Thank you for participating in this survey. Click the "next" button to be redirected to a separate page where you can anonymously provide your name, and instructor's name, to receive bonus points for participation.

APPENDIX E: Product Webpage Stimuli (Study Three)

4 out of 5 (High Volume)

4 out of 5 (Medium Volume)

4 out of 5 (Low Volume)

8 out of 10 (High Volume)

8 out of 10 (Medium Volume)

All-new
isoTech Tab Pro HD8
12-hour battery | 2X the storage

isoTech Tab Pro Tablet, 8'' HD Display, 16GB, Wifi, Bluetooth 4.0 (Black)
by LabTech

★★★★★★★★☆☆ 90 customer reviews

Price: $89.99 & FREE Shipping.

Color: Black

- 1.6 GHz Quad-Core Processor
- Supports Wi-Fi connection
- Runs Android 5.1 (Lollipop) OS
- 1.5 GB RAM memory
- 8-inch HD display
- Screen resolution of 1280 x 800 pixels
- 16 GB of internal storage
- Micro SD slot for up to 200 GB of expandable storage
- Up to 12 hours of battery life
- 2 MP front camera and 5 MP rear camera

---

8 out of 10 (Low Volume)

All-new
isoTech Tab Pro HD8
12-hour battery | 2X the storage

isoTech Tab Pro Tablet, 8'' HD Display, 16GB, Wifi, Bluetooth 4.0 (Black)
by LabTech

★★★★★★★★☆☆ 15 customer reviews

Price: $89.99 & FREE Shipping.

Color: Black

- 1.6 GHz Quad-Core Processor
- Supports Wi-Fi connection
- Runs Android 5.1 (Lollipop) OS
- 1.5 GB RAM memory
- 8-inch HD display
- Screen resolution of 1280 x 800 pixels
- 16 GB of internal storage
- Micro SD slot for up to 200 GB of expandable storage
- Up to 12 hours of battery life
- 2 MP front camera and 5 MP rear camera

APPENDIX F: Complete Stimulus for All Experimental Conditions (Study Three)

Research Information Sheet

Title of Study: Star Rating Evaluation Survey

Principal Investigator (PI): Aaron Johnson (Marketing Department) 313-577-4406

Purpose: You are being asked to be in a research study of online products presented in a product webpage because you are a potential consumer of this type of product. Study Procedures: If you take part in the study, you will be asked to:

* Fill out a survey.

* Answer questions about your attitude toward the ad that you will view and also your online shopping experiences. In addition, you will be asked background questions on your demographic information. You may skip background related questions and still complete the survey.

* This is a one-time survey that should take approximately 10 minutes to complete.

Benefits: As a participant in this research study, there will be no direct benefit for you; however, information from this study may benefit other people now or in the future.

Risks: There are no known risks at this time to participation in this study.

Costs: There will be no costs to you for participation in this research study.

Compensation: For taking part in this research study, you will be paid for your time via bonus points awarded by your instructor.

Confidentiality: All information collected about you during the course of this study will be kept without any identifiers.

Voluntary Participation /Withdrawal: Taking part in this study is voluntary. You are free to withdraw at any time. Your decision will not change any present or future relationships with Wayne State University or its affiliates.

Questions: If you have any questions about this study now or in the future, you may contact Aaron Johnson at the following phone number 313-577-4406. If you have questions or concerns about your rights as a research participant, the Chair of the Institutional Review Board can be contacted at (313) 577-1628. If you are unable to contact the research staff, or if you want to talk to someone other than the research staff, you may also call the Wayne State Research Subject Advocate at (313) 577-1628 to discuss problems, obtain information, or offer input.

Participation: By completing the questionnaire you are agreeing to participate in this study. Click the "Next" button below to continue.

This survey will first have you respond to statements about an electronic tablet. Next, you will view an ad for the electronic tablet and answer the subsequent questions after each ad. There are no right or wrong answers, but we ask that you please take your time and answer truthfully. If you agree to participate, please click the button below to begin the survey.

Please indicate how much you agree or disagree with the following statements regarding electronic tablets:

I know pretty much about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Compared to most other people, I know less about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I am very knowledgeable about the product category of tablets.
(1-Strongly Disagree; 9-Strongly Agree)

I do not feel very knowledgeable about tablets.
(1-Strongly Disagree; 9-Strongly Agree)

Please carefully review the product webpage below and then answer the questions on the following pages.

According to the product webpage, what was the star rating of the product?
(1-4 out of 5 stars; 2-8 out of 10 stars)

According to the product webpage, what was the number of consumer ratings for the product?
(1-15; 2-90; 3-1394)

The quality of this product seems to be
(1-Very Low; 9-Very High)

This product seems to be reliable.
(1-Strongly Disagree; 9-Strongly Agree)

The manufacturing quality of this product seems to be
(1-Very Low; 9-Very High)

This product seems to be dependable.
(1-Strongly Disagree; 9-Strongly Agree)

This product is likely to be durable.
(1-Strongly Disagree; 9-Strongly Agree)

If you were going to buy a tablet, how likely would you be to buy the tablet shown in the product webpage?
(1-Not Likely at All; 2-Very Likely)

Given the information in the product webpage, the likelihood of purchasing this product is
(1-Very Low; 9-Very High)

How probable is it that you would consider purchasing this product?
(1-Not Probable; 2-Very Probable)

The next few questions ask about your responses to the product webpgae you saw at the start of the survey.

How familiar are you with the isoTech® brand of tablets?
(1-Not at All Familiar; 9-Very Familiar)

How believable is the product webpage?
(1-Not at All Believable; 9-Highly Believable)
(1-Not at All Acceptable; 9-Highly Acceptable)

What is your overall attitude toward the product webpage?
(1-Unfavorable; 9-Favorable)
(1-Bad; 9-Good)
(1-Negative; 9-Positive)

Select 8 on the scale below to demonstrate that you are reading carefully.
(1-Strongly Disagree; 9-Strongly Agree)

The price shown for the product is
(1-Very Unacceptable; 2-Very Acceptable)

This product is a
(1-Very Poor Value for the Money; 2-Very Good Value for the Money)

The 5-star rating scale in the product webpage is commonly used. <u>OR</u> The 10-star rating scale in the product webpage is commonly used.
(1-Strongly disagree; 9-Strongly Agree)

Please report the number of times you have seen a 5-star rating scale used for online product ratings. OR Please report the number of times you have seen a 10-star rating scale used for online product ratings.
(1-Never Seen Before; 2-Seen Many Times)

I am familiar with the 5-star rating scale used for online product ratings. OR I am familiar with the 10-star rating scale used for online product ratings.
(1-Strongly disagree; 9-Strongly Agree)

I feel comfortable using the 5-star rating scale to guide my online purchase decisions. OR I feel comfortable using the 10-star rating scale to guide my online purchase decisions.
(1-Strongly disagree; 9-Strongly Agree)

Please answer the following two questions WITHOUT returning to the product webpage you viewed previously:

Please report the star rating (number of yellow-filled stars) for the product on the webpage you viewed.
(Text entry)

Please report the number of customer who have provided a rating for the product on the webpage you viewed.
(Text entry)

When I buy a product online, I always check the ratings that are presented on the website.
(1-Strongly disagree; 9-Strongly Agree)

When I buy a product online, the ratings presented on the website are helpful for my decision-making.
(1-Strongly disagree; 9-Strongly Agree)

When making online purchases, I trust the consumer star ratings to guide my decisions.
(1-Strongly disagree; 9-Strongly Agree)

When I buy a product online, the ratings presented on the website make me confident in purchasing the product.
(1-Strongly disagree; 9-Strongly Agree)

I am an experienced online shopper.
(1-Strongly disagree; 9-Strongly Agree)

On average, how often do you make online purchases each month?
(______________________________)

Do you currently own an electronic tablet?
(1-Yes; 2-No)

Are you currently looking to purchase an electronic tablet?
(1-Yes; 2-No)

What operating system platform would you prefer to use in an electronic tablet?
(1-Apple iOS; 2-Android; 3-Other)

What is your gender?
(1-Male; 2-Female)

What is your age?
(______________)

Thank you for participating in this survey. Click the "next" button to be redirected to a separate page where you can anonymously provide your name, and instructor's name, to receive bonus points for participation.

REFERENCES

Alter, A. L., & Oppenheimer, D. M. (2008). Easy on the mind, easy on the wallet: The roles of familiarity and processing fluency in valuation judgments. *Psychonomic Bulletin & Review*, *15*(5), 985-990.

Bagchi, R., & Davis, D. F. (2012). $29 for 70 items or 70 items for $29? How presentation order affects package perceptions. *Journal of Consumer Research*, *39*(1), 62-73.

Biswas, A., Bhowmick, S., Guha, A., & Grewal, D. (2013). Consumer evaluations of sale prices: role of the subtraction principle. *Journal of Marketing*, *77*(4), 49-66.

Bonato, M., Fabbri, S., Umiltà, C., & Zorzi, M. (2007). The mental representation of numerical fractions: Real or integer?. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1410.Chen, P.-y., et al. (2006). "All reviews are not created equal."

Chen, Y. B., et al. (2011). "Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning." *Journal of Marketing Research* 48(2): 238-254.

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

Chintagunta, P. K., et al. (2010). "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets." *Marketing Science* 29(5): 944-957.

Churchill Jr, G. A., & Peter, J. P. (1984). Research design effects on the reliability of rating scales: A meta-analysis. *Journal of Marketing Research*, 360-375.

Clemons, E. K. and G. D. Gao (2008). "Consumer informedness and diverse consumer purchasing behaviors: Traditional mass-market, trading down, and trading out into the long tail." *Electronic Commerce Research and Applications* 7(1): 3-17.

Clemons, E. K., et al. (2006). "When online reviews meet hyperdifferentiation: A study of the craft beer industry." *Journal of Management Information Systems* 23(2): 149-171.

Coulter, K. S., & Coulter, R. A. (2007). Distortion of price discount perceptions: The right digit effect. *Journal of Consumer Research*, *34*(2), 162-173.

Decker, R. and M. Trusov (2010). "Estimating aggregate consumer preferences from online product reviews." *International Journal of Research in Marketing* 27(4): 293-307.

Dellarocas, C., et al. (2007). "Exploring the value of online product reviews in forecasting sales: The case of motion pictures." *Journal of Interactive Marketing* 21(4): 23-45.

Duan, W., et al. (2008). "Do online reviews matter?—An empirical investigation of panel data." *Decision Support Systems* 45(4): 1007-1016.

Duan, W., et al. (2008). "The dynamics of online word-of-mouth and product sales - An empirical investigation of the movie industry." *Journal of Retailing* 84(2): 233-242.

Epley, N., & Gilovich, T. (2010). Anchoring unbound. *Journal of Consumer Psychology*, *20*(1), 20-24.

Fischer, M. H., & Fias, M. H. (2005). Spatial representation of numbers. *Handbook of Mathematical Cognition*, 43-54.

Flanagin, A. J. and M. J. Metzger (2013). "Trusting expert- versus user-generated ratings online: The role of information volume, valence, and consumer characteristics." *Computers in Human Behavior* 29(4): 1626-1634.

Garner, W. R., & Hake, H. W. (1951). The amount of information in absolute judgments. *Psychological Review*, *58*(6), 446.

Ganu, G., et al. (2013). "Improving the quality of predictions using textual information in online user reviews." *Information Systems* 38(1): 1-15.

He, S. X. and S. D. Bond (2015). "Why is the crowd divided? Attribution for dispersion in online word of mouth." *Journal of Consumer Research* 41(6): 1509-1527.

Heintz, R. K. (1950). The effect of remote anchoring points upon the judgment of lifted weights. *Journal of Experimental Psychology*, *40*(5), 584.

Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review*, *55*(6), 297.

Ho-Dac, N. N., et al. (2013). "The effects of positive and negative online customer reviews: do brand strength and category maturity matter?" *Journal of Marketing* 77(6): 37-53.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). Communication and persuasion; psychological studies of opinion change.

Ischebeck, A., Schocke, M., & Delazer, M. (2009). The processing and representation of fractions within the brain: An fMRI investigation. *NeuroImage*, *47*(1), 403-413.

Janiszewski, C., & Lichtenstein, D. R. (1999). A range theory account of price perception. *Journal of Consumer Research*, *25*(4), 353-368.

Khare, A., et al. (2011). "The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings." *Journal of Retailing* 87(1): 111-126.

Koh, N. S., et al. (2010). "Do online reviews reflect a product’s true perceived product quality? An investigation of online movie reviews across cultures." *Electronic Commerce Research and Applications* 9(5): 374-385.

Komorita, S. S., & Graham, W. K. (1965). Number of scale points and the reliability of scales. *Educational and Psychological Measurement*, *25*(4), 987-995.

Li, X. X. and L. M. Hitt (2008). "Self-Selection and Information Role of Online Product Reviews." *Information Systems Research* 19(4): 456-474.

Moe, W. W. and D. A. Schweidel (2012). "Online Product Opinions: Incidence, Evaluation, and Evolution." *Marketing Science* 31(3): 372-386.

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

Mousavizadeh, M., et al. (2015). "The Effect of Central and Peripheral Cues on Online Review Helpfulness: A Comparison between Functional and Expressive Products." *Conference on Information Systems*.

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

Nunnally, J. C., & Bernstein, I. H. (1978). Psychometric theory.

Ostrom, T. M., & Upshaw, H. S. (1968). Psychological perspective and attitude change. *Psychological Foundations of Attitudes*, 217-242.

Park, C. W., & Lessig, V. P. (1981). Familiarity and its impact on consumer decision biases and heuristics. *Journal of Consumer Research*, *8*(2), 223-230.

Park, Whan C., Meryl P. Gardner and Vinod K. Thukral (1988), “Self-Perceived Knowledge: Some Effects on Information Processing for a Choice-Task,” *American Journal of Psychology*, 101 (Fall), 401–24.

Park, D.-H. and S. Kim (2009). "The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews." *Electronic Commerce Research and Applications* 7(4): 399-410.

Park, D.-H. and J. Lee (2009). "eWOM overload and its effect on consumer behavioral intention depending on consumer involvement." *Electronic Commerce Research and Applications* 7(4): 386-398.

Schlosser, A. E. (2011). "Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments." *Journal of Consumer Psychology* 21(3): 226-239.

Sen, S. and D. Lerman (2007). "Why are you telling me this? An examination into negative consumer reviews on the web." *Journal of Interactive Marketing* 21(4): 76-94.

Sherif, M., & Hovland, C. I. (1961). Social judgment: Assimilation and contrast effects in communication and attitude change.

Sun, M. (2012). "How does the variance of product ratings matter?" *Management Science* 58(4): 696-707.

Tsang, A. S. and G. Prendergast (2009). "Is a “star” worth a thousand words? The interplay between product-review texts and rating valences." *European Journal of Marketing* 43(11/12): 1269-1280.

Townsend, C., & Kahn, B. E. (2013). The “visual preference heuristic”: The influence of visual versus verbal depiction on assortment processing, perceived variety, and choice overload. *Journal of Consumer Research*, *40*(5), 993-1015.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

Ullah, R., et al. (2016). "From valence to emotions: Exploring the distribution of emotions in online product reviews." *Decision Support Systems* 81: 41-53.

Viswanathan, M., Sudman, S., & Johnson, M. (2004). Maximum versus meaningful discrimination in scale response:: Implications for validity of measurement of consumer perceptions about products. *Journal of Business Research*, *57*(2), 108-124.

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. *The psychology of evaluation: Affective processes in cognition and emotion*, 189-217.

Wu, P. F. (2013). "In Search of Negativity Bias: An Empirical Study of Perceived Helpfulness of Online Reviews." *Psychology & Marketing* 30(11): 971-984.

Ye, Q., et al. (2009). "The impact of online user reviews on hotel room sales." *International Journal of Hospitality Management* 28(1): 180-182.

Zhang, J. Q., et al. (2010). "When does electronic word-of-mouth matter? A study of consumer product reviews." *Journal of Business Research* 63(12): 1336-1341.

Zhang, L., et al. (2013). "The impact of online user reviews on cameras sales." *European Journal of Marketing* 47(7): 1115-1128.

Zhu, F. and X. Q. Zhang (2010). "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics." *Journal of Marketing* 74(2): 133-148.

Zou, P., et al. (2011). "Does the valence of online consumer reviews matter for consumer decision making? The moderating role of consumer expertise." *Journal of Computers* 6(3): 484-488.

ABSTRACT

STAR-CROSSED CONSUMERS: THE EFFECTS OF ONLINE RATING SCALE LENGTH ON PRODUCT EVALUATIONS

by

AARON JOHNSON

August 2017

Advisors: Dr. Abhijit Biswas and Dr. Sujay Dutta

Major: Business Administration (Marketing)

Degree: Doctor of Philosophy

Consumers' ratings of products are ubiquitous in the online marketplace (e.g., Amazon; Yelp). The rating scales provided by online businesses typically comprise a set of stars that appear in the form of linear scales. Consumers looking to purchase a certain product likely rely on product ratings based on these rating scales. Although past research confirms the intuitive expectation that a higher star rating for a product elicits more favorable responses from consumers, there is a paucity of research related to effects of the properties of the scales themselves on consumers' psychology. The literature on cognitive processing of information suggests that varying properties of scales might affect people's processing of them and in turn their perceptions. Both 5-point and 10-point star-based rating scales, i.e., scales with a total of 5 and 10 stars respectively, are common in the online marketplace. Using relevant theories from the cognitive processing literature, this dissertation investigates whether the number of scale points in a rating scale affects consumers' perceptions of product quality and their purchase intention. The results of three studies show that when a specific rating (e.g., 80%) is presented on a 10-point star-based scale (i.e., 8 out of 10 stars), perceptions of product quality and consumers' intention to purchase the product are higher compared to when the same rating is presented on a

5-point scale (i.e., 4 out of 5 stars). Implications and limitations of this research are discussed, and directions for further research are provided.

AUTOBIOGRAPHICAL STATEMENT

Aaron Johnson began his bachelor's degree at the University of California, Davis, and completed it at Brigham Young University in 2008. He then earned a master's degree in business administration in 2013 from the Woodbury School of Business, at Utah Valley University. While completing his undergraduate and graduate education, Aaron started a family with his wife and they welcomed two daughters into their lives. Concurrent to his schooling, Aaron gained 10 years of full-time industry experience in sales, marketing, and relationship management before joining the marketing doctoral program in 2013 in the Mike Ilitch School of Business, at Wayne State University. As a graduate student, Aaron began to develop a love for research and was fortunate to work on a project which was later published in a marketing journal. As a doctoral student, Aaron presented several papers at national marketing conferences and has taught Principles of Marketing, Business Statistics, and Personal Selling and Sales Management at Wayne State University. His primary research interest is consumer perceptions of marketplace signals, particularly online consumer ratings. Aaron will receive the Doctor of Philosophy degree in Business Administration (with concentration in marketing) in August 2017.