CONTENT AND CONTEXT: CONSUMER INTERACTIONS WITH DIGITAL DECISION AIDS

A Dissertation submitted to the Temple University Graduate Board

---

In Partial Fulfillment of the Requirements for the Degree DOCTOR OF PHILOSOPHY OF BUSINESS ADMINISTRATION

By Patrick A. Barbro

Diploma Date July 2015

Examining Committee Members:

Susan M. Mudambi, Advisory Chair, Department of Marketing and Supply Chain Management
Eric M. Eisenstein, Department of Marketing and Supply Chain Management
Robert J. Kent, External Member, University of Delaware
Maureen Morrin, Department of Marketing and Supply Chain Management
David Schuff, Department of Management Information Systems

ABSTRACT

Through four essays, this dissertation contributes to the body of marketing literature by advancing understanding of consumer interactions with digital decision aids. Different aspects of the content contained within digital decision aids are explored in several contexts. First, the drivers of consumer interactivity in an online review community are examined and it is found that violations of community norms are an important factor in stimulating consumer action. Second, a tool is developed to facilitate the normalization of online review content across languages. Next, elements of language and national culture are investigated to determine their influence on consumer reviews in an international context. It is found that cultural biases play an important role in the relative verbosity, valence, and helpfulness of online reviews across countries. Lastly, the role of images in digital decision aids is considered and it is found that image type and perspective can influence consumer product evaluation. In sum, the influence that content and context have on consumer interactions with digital decision aids is clearly demonstrated through a diverse yet intertwined set of studies.

This dissertation is dedicated to my beautiful wife Mary Beth. Without her kindness, help, love, and support, this would never have been possible. I hope that we can share in this accomplishment and those that lay ahead of us for years to come.

ACKNOWLEDGMENTS

I would like to thank my committee members Eric Eisenstein, Bob Kent, Mimi Morrin, and David Schuff for their continued support and and assistance throughout this process. Special thanks to my Advisor, Susan Mudambi, for her gracious handling of the years of advice and patiently dealing with me throughout. Who would have thought that volunteering for some data collection five years ago would have lead us on this adventure all the way here. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The quantity and variety of information available to assist consumers with purchase decisions is overwhelming. Current technology gives consumers 24/7 access to millions of pieces of potentially influential information at their fingertips. The structure of this information can take the form of tools that are designed and provided specifically to assist in consumer decision making. The format of these widespread digital consumer decision aids can take on many forms. In stores, consumers have accesses to tools like self-service kiosks and custom fit technologies. Tools like online reviews are available to consumers at any hour while shopping online at home, but can also be accessed via mobile devices while evaluating products in stores. The boundary between online and offline decision aids is becoming increasingly blurred by by the power and mobility of the smartphones carried by many shoppers. Technologies like virtual try-on for example allow consumers to "try-on" products anytime, anywhere. With this incredible level of access to information that may guide purchase decisions, how consumers perceive, use, and interact with these decision aids is critical to marketers.

This dissertation will add to the body of marketing knowledge through four essays examining factors concerning how consumers interact with digital decision aids. More specifically, this research will consider the impact of specific elements of the content present within or delivered by these decision aids and the context in which consumers encounter them. When examining content contained within digital decision aids, there are several different aspects of the content that can be considered. In the case of online reviews, elements of the written text such as length, valence, helpfulness, language, and meaning behind the written words are key elements of the review content that need to be considered. Images and aspects of them are another type of content that is important both to online reviews and other types of digital decision aids like virtual try-on. In terms of context for online reviews, norms of the online review community, where they are placed on a website, and the national culture of the shopper are key contextual elements to consider. Beyond online reviews considering where consumers interact with decision aids, and the personal habits of the individuals using them are key aspects of context to consider. By examining gaps in online review literature and addressing the influence of newer decision aid technology, this dissertation will add to marketing theory and practice by providing a better understanding of how these elements of content and context influence how consumers interact with these important shopping tools.

Online reviews are of particular importance due to their widespread availability and use. For example, the well-known review site Yelp! alone has over 77 million consumer reviews of local businesses. Online ratings from sites like Yelp! have been shown to have a substantial impact on firm revenues making them of prime importance to marketers (Luca 2011). While online reviews have been widely adopted and researched, continually advancing technology has allowed for the development of new devices that can aid consumer decision making as well (Varadarajan et al. 2010). Virtual Try-on is one type of technology that allows consumers to create virtual models of themselves to "try-on" apparel when shopping online. Virtual try-on technologies like these have been shown to increase the entertainment value of the online shopping experience (Kim and Forsythe 2008). Regardless of the format these shopping tools may take, the content contained within them, and context of how they are perceived are important considerations for the shoppers that use them, the managers that deploy them, and the researchers that study them.

One of the most studied sources of electronic information used by consumers to aid in shopping decisions is consumer generated online reviews (e.g. Chevalier and Mayzlin 2006; Zhu and Zhang 2010). When considering online reviews, content and context are both multifaceted. Elements of review content include: opinions offered by the reviewer through text, star ratings provided, nature of the language used, images, and additional content added on by other consumers. Elements of review context include: where a review is placed on a website, the norms of the community they review exists in, and even the country in which the review was written.

Much of the prior research concerning online reviews has considered their impact on sales and the influence of specific review characteristics. Consumer generated online reviews have been shown to influence sales in several contexts highlighting the importance of this topic to both academics and managers (Chevalier and Mayzlin 2006; Zhu and Zhang 2010). Chevalier and Mayzlin (2006) for example have shown that improved book reviews can lead to an increase in sales. Extending the influence of reviews on sales to consider certain review characteristics, several researchers have found that volume, but not valence leads to improved product sales (Chen, Wu, and Yoon 2004; Liu 2006; Duan, Gu, and Whinston 2008). Valence and helpfulness have shown to be important characteristics of online review content in the literature. Sen and Lerman (2007) for example show that negative reviews garner more attention from consumers, demonstrating that elements of review content can stimulate consumer consideration. Review characteristics like valence can also draw a consumer response in terms of how helpful they find the content (Mudambi and Schuff 2010). Within the online review research examining valence, the variance in consumer ratings has also been shown to be an important factor (Sun 2012). Further research has examined social dynamics associated with online reviews and demonstrated that the environment or context is an influential factor in consumer reviewing behavior. The influence of other consumers has been shown to impact product ratings, sales, and product evaluation (Sridhar and Srinivasan 2012; Moe and Schweidel 2011; Moe and Trusov 2011). Written review content and style have also been shown to impact conversion rates and helps to determine what reviews are most influential (Ludwig et al. 2013). It is also worth noting that many studies concerning online reviews are cross-sectional in nature, making it difficult to assess how the review environment influences consumers as it changes over time.

Online reviews have become more than just a source of information for consumers, but an interaction point where consumer to consumer communication influences both the reviews and the perception of them. It stands out that research concerning how consumers actually interact with online reviews and why is more limited and underrepresented in the literature. Chapter one of this dissertation employs a novel dataset to evaluate drivers of consumer interaction with online reviews when products are first released. Online reviews provide an opportunity for consumers to engage with opinions provided by their peers, thereby creating an interactive forum. The interactions that occur shortly after a product is released can have consequences long after they happen. Many retailers and review sites provide the opportunity for consumers to comment on the reviews of others or “vote” as to whether they think the review is helpful, funny, or cool. Comments on reviews add additional content for consumers to consider, while voting systems such as "helpfulness" can alter which specific reviews consumers are first presented with on popular retailing sites like Amazon.com.

While past research has helped explain the impact of social dynamics on future rating behavior (eg. Moe and Schweidel 2011; Moe and Trusov 2011; Sridhar and Srinivasan 2012), this research examines the influence of these dynamics on community interactions with existing reviews immediately after a product launches. Three key elements are considered when evaluating the dynamics of consumer interactions with reviews at product launch: community norms, visual accessibility, and social dynamics. By evaluating how the early dynamics of the review environment in conjunction with review characteristics and website placement can lead to customer interactions, this study adds to the conversation with theoretical, methodological, and managerial contributions.

To accomplish this, both the content of the reviews and the context in which consumers perceive them must be considered. Chapter one evaluates content of reviews through structural elements such as valence and affective language, but also considers how reviews fit within the context of the expected norms of the reviewing community. An evaluation of where and when consumers encounter a review and prior consumer interactions with them will provide contextual elements critical to this phenomenon. By examining these factors together, this study develops for a greater understanding of the drivers behind consumer interactions with online reviews at product launch. The findings of chapter one demonstrate that violations of online community norms, website placement, and prior interactions from other consumers are influential factors present in the content and context of online reviews that generate consumer interactions with them. Employing a unique dataset that tracks placement, content, and interaction with specific reviews over time enables the impact of each of these variables to be accounted for, and offers a methodological contribution to the literature as well.

How online reviews differ cross-culturally has been identified as an area that is underrepresented in the literature (King, Racherla, and Bush 2014). Online review content is now generated in a myriad of different languages in the context of country specific retail sites of both local and multinational retailers. Lacking in the literature is a cross-national view of how country and language differences may influence the use and characteristics of online reviews. Amazon for example has 14 country specific sites covering nine different languages and it is estimated that only 55% of web content is in English (W3Techs 2014). Despite the proliferation of online review content across borders, there is little academic research examining the influence of online review content in an international context. To address these gaps in the literature, a cross-cultural analysis of the influence of review content and the context they are available to consumers is needed. Chapters three and four of this dissertation address this gap by examining the influence of language and culture consumer interactions with online reviews.

Chapter three examines some of the most basic building blocks of online review content. Language is the basis of any written content generated by consumers, but the influence of language on online reviews has been largely overlooked in the literature. The quantity of web chatter has been linked to product sales and firm performance while longer online reviews have been shown to be more diagnostic and helpful to consumers (Dhar and Chang 2009; Mudambi and Schuff 2010; Tirunillai and Tellis 2012). These findings do not take into account however that the amount of content provided in a review of a given length can greatly vary depending on the language it is written in. Chapter three establishes that language is an important factor to consider when comparing the volume of online review content across countries.

Firms care about the length of online reviews as it can represent consumer interest, engagement, and involvement (Bughin et al. 2010; Chen et al. 2011; Liu 2006). To compare review length across languages however, multinational retailers need analysis methods that are scalable and meaningful. Despite the flood of online content around the world, there is no generally accepted method to compare the verbosity of consumer word of mouth across languages. Chapter three offers a solution to this issue through the development of an index that allows for the comparison of online reviews across languages. This index, called the Character Count Language Index (CCLI), removes the structural language differences that account for variation in the length of review content across countries. The CCLI provides a tool that by accounting for native language, normalizes the length of review content across languages. This normalization facilitates better comparisons of consumer word of mouth from country to country, allowing for more accurate and meaningful analysis by mangers and researchers.

In addition to the structural language differences present in reviews across countries, the culture and norms present in the countries where reviews are written provide a basic context for how a review may be constructed and received by consumers. Chapter four further addresses the language and culture gap in online review research. By examining a large dataset of reviews from the United States, the United Kingdom, Germany, France, and Japan, chapter four begins to establish how culture and language influence the verbosity, valence, and helpfulness of online reviews across countries.

These insights will be of particular interest to multinational retailers who use country based websites.

First, chapter four applies the Character Count Language Index developed in chapter three in conjunction with Hall’s (1976) high-context low-context categorization of communication styles to establish that both national culture and language structure are influential in determining the relative verbosity of review content across countries. Next, cultural response biases are considered in an analysis of the valence of reviews across countries. It is found that the culture of the country in which a review is written can impact the aggregate star rating of reviews. Finally, culture is also considered as a factor in determining the relative helpfulness of reviews across countries. The results show that culture plays a role in the perceived helpfulness of reviews in different countries. The results also suggest that prior research analyzing review helpfulness may not be applicable across different countries, cultures, and languages. Overall, the results found in chapter four establish that practitioners and researchers must take language and culture into account when evaluating consumer review content in the context of a multinational setting.

Consumer generated content in digital decision aids extends beyond written text and star ratings. While sitting in a restaurant it is common to see patrons snapping pictures of their food, the scenery, and themselves. Many of these images end up in online reviews as can be confirmed by scrolling through restaurant reviews on Yelp! This imagery adds to the traditional text evaluation and star ratings that are included in online reviews to create a richer review and evaluation of the consumer experience. The use of consumer pictures extends well beyond online reviews however. Modern technology is now allowing the use of digital decision aids that are more mobile, sharable, and seamless with normal shopping behavior, creating different contexts for consumer evaluation. The ease with which consumers can capture digital images of themselves via ‘selfies’ even facilitates their ability to “try on” products in a wider variety of environments. One manifestation of this, called virtual mirrors, serves as an example. Using personal electronic devices, virtual mirrors enable consumers to view themselves using or trying on a product without physically examining it in person. The type of self-image captured while using virtual mirrors, and how the customer perceives it, are likely to play a role in consumers’ evaluation of the products and product images. Technologies like these illustrate the importance for marketers and researchers to better understand how consumers use and perceive images that may influence product evaluation and purchase decisions.

The proliferation of smartphones, social media, selfies, and photo sharing has changed the way consumers have become used to seeing themselves. It is generally understood that people generally prefer mirror image (i.e., what they see when they look in the mirror) to their true image (i.e., how others see them), because it is more familiar to them (Mita et al. 1977). Researchers have begun to examine this issue in the context of digital decision aids. Cho and Schwarz (2012) for example look at image type in the context of virtual mirrors and found that while people liked their mirror image more than their true image, products shown on their mirror image did not receive an evaluation advantage. Due to advances in technology, decision aids, and the frequency in which consumers take and see themselves in pictures, traditional assumptions about consumer image preference need to be revisited. Chapter five of this dissertation takes this step by evaluating how consumer selfie taking practices can influence their preference for certain types of images and how they evaluate products in them.

Through three experiments, Chapter five establishes a “selfie effect” which shows that consumers who frequently take selfies are generally more likely to positively evaluate products seen in multiple types of images. This effect also shows that consumers who frequently take selfies more strongly favor their true image when evaluating products. The findings of chapter five challenge the traditional assumptions about consumer image preference and show that marketers and academics need to acknowledge that modern media, technology, and consumer habits are changing the way consumers view themselves. These findings also have important consequences for consumer interactions with digitial decision aids including online reviews and virtual try-on applications. By better understanding how consumers view themselves and products in images, marketers can better design and interpret consumer uses of them in the context of digital decision aids.

In whole, the four essays of this dissertation add to the body of marketing knowledge by analyzing under researched aspects of consumer interactions with digital decision aids. By focusing on several specific elements of the content and context in which these interactions occur, the findings add clarity to consumer attitudes, decisions, and behavior and how managers and researchers can best interpret them.

CHAPTER 2: ONLINE CONSUMER-REVIEW INTERACTIONS: WHAT'S THE NORM?

Introduction

Using online reviews as a valuable information resource is commonplace for shoppers. The process has become more than reading a review, integrating the information, and proceeding with the purchase decision. Shoppers also face the option of interacting with reviews in other ways. They are asked to vote whether a review is helpful (in the case of Amazon), funny or cool (in the case of Yelp!), and to provide and respond to reviews and their comments. This makes the online review a highly interactive "touch point" that allows consumers to interact with both reviews and each other. This interactivity can convey a vibrant online community that encourages additional site engagement. If there is little review interactivity, particularly when a product is first released, it could serve as a negative signal to other consumers about the product. It is well-understood that utilizing online customer reviews is a valuable exercise for consumers (King, Racherla, and Bush 2014), but it is less clear which factors encourage high review interactivity when a product is first released. This paper investigates elements of the review, the reviewer, and the review environment that are contributing factors to interactivity, and examines how the continually changing nature of these factors affects interactions with reviews over time.

This research addresses these questions by building on the work of Moe and Trusov (2011), Moe and Schweidel (2011), and Sridhar and Srinivasan (2012), who establish that social dynamics of the review environment impact product ratings, the volume and characteristics of future reviews, and future sales. While past research has helped explain the impact of social dynamics on future rating behavior, this study examines the influence of these dynamics on community interactions with existing content. By evaluating how the early dynamics of the review environment in conjunction with review characteristics and website placement lead to interactivity, this study adds to the theoretical, methodological, and managerial body of knowledge.

Moe and Schweidel (2011) find that the rating environment impacts an individual's decision whether or not to post a review. More specifically, they show positive review environments can increase the incidence of posting reviews. Sridhar and Srinivasan (2012) find social influences impact how consumers evaluate and rate products. These findings demonstrate the important social dynamics within online review communities, and point to additional impacts that need to be explored.

Moe and Trusov (2011) conceptualize social dynamics as "the impact of previously posted ratings on the posting of future ratings" (p. 447). This conceptualization works well for the broad goal of determining the impact of dynamics on sales in an online review environment. To better understand the potential impact of the review environment on consumer behavior in online review communities, research can examine these impacts at a finer, more micro level. Examining individual reviews as they are posted allows for an understanding of how specific characteristics of a review can influence the actions of other consumers.

Adherence to norms is one factor in these social dynamics. Past research has established online review forums as a specific type of online community (Forman, Ghose, and Wiesenfeld 2008) with its own norms and characteristics (Moon and Sproull 2008;

Ren, Kraut, and Kiesler 2007). This study examines consumer reactions to social norms present in the online review community as a driver of interactivity with online reviews.

Another influence on social dynamics is placement of the review on a site's page. Retailers strategically place certain reviews in more visually prominent locations, such as the product landing page. Ludwig et al. (2013) suggest that it is important for managers to consider how reviews are placed and promoted on their sites because features of certain reviews, such as affective content, may improve conversion rates. Many online retailers determine the placement of a review by the level and type of prior consumer interaction with that review. This means that how consumers respond to a particular review can influence what reviews future shoppers will see first. A review can get "stuck" in a prominent position on a retailer's product page if it sustains high levels of consumer response. Depending on the content of that review, this could have beneficial or painful consequences for the sellers of that product.

Visually prominent book reviews create a "digital book jacket" that can accompany a product for hours, days, or months on the retail site. A physical book jacket displays positive, enticing reviews from famous authors and literary experts. However, Amazon prominently displays both the most helpful positive review and the most helpful negative review. Consumers go to review sites to look for useful information, not for sales pitches, and potentially benefit when reviews are not universally positive. Since information printed on the digital book jacket can influence consumer attitudes and decisions, it is important to understand the effect visual prominence has on reviews, and how these prominent placements influence consumer/review interactivity.

Methodologically, the unique data collection technique used in this study presents a new way to study online reviews by observing consumer interaction with the reviews over time in a continually changing marketplace, instead of relying on a snapshot view. The unique data collection method allows for analysis of how changes in the Amazon.com review environment and social dynamics interact with review characteristics, and their combined impact on customer interactions with the reviews as they happens. Results show that the factors work in concert to influence customer interaction with online reviews. By broadening the discussion of social dynamics to include customer interactions with online reviews, this research adds to prior theory and helps explain an important phenomenon.

The balance of this paper is structured as follows. The next section provides a brief review highlighting the key contributions in research concerning online reviews. This is followed by the conceptual development of hypotheses considering: community norms, visual prominence, and social dynamics. Next, the data, methods, and results are presented. The paper concludes with research limitations and a discussion of contributions to theory and practice.

Literature Review

Online communities are contexts in which consumers can influence other consumers' options on products, brands, and services (Kozinets 2002). Forman, Ghose, and Wiesenfeld (2008) refer to online communities as "voluntary collectives whose members share a common interest or experience and who interact with one another primarily over the internet" (p.293). Online review forums, such as Amazon.com, represent a specific type of online community where some consumers share their opinions and experiences about products and services and others consume those reviews as information for a future purchase decision (Chevalier and Mayzlin 2006). These forums also allow for consumer interactions with reviews themselves. Research on online reviews can broadly be categorized into three streams: impact of reviews on sales and marketing strategy; influence and consequences of review characteristics; and most consequential to this research, dynamics of consumer rating behavior.

Effects of Reviews on Sales and Strategy

Chevalier and Mayzlin (2006) show that improved book reviews led to a relative increase in sales on a given website, demonstrating the importance of online reviews in driving sales. This study draws attention to the importance of review valence in two ways. Higher star ratings are shown to be associated with higher sales, demonstrating the influence of consumer opinion on sales. Review extremity is also shown to have an important effect, as one star reviews were highly influential. On both Amazon.com and bn.com, they also find far fewer negative reviews than positive reviews.

The effect of reviews on sales has not always been consistent, however. Chen, Wu, and Yoon (2004) find that while the quantity of consumer book reviews on Amazon.com is positively associated with sales, consumer ratings are not. The inconsistency of results concerning the impacts of consumer ratings on sales extends to movies. Liu (2006) shows that word of mouth found online is positively associated with box office revenue, but only in terms of quantity, not valence. In one of only a few studies using a longitudinal sample, Duan, Gu, and Whinston (2008) find that volume, but not valence, is related to increased sales. Zhu and Zhang (2010) incorporated product and consumer characteristics to show other aspects of the environment can influence the effect of online reviews on sales. Berger, Sorensen, and Rasmussen (2010) studied product awareness and find, relative to not being reviewed at all, even a negative review can have a positive effect on book sales.

Influence of Review Characteristics

Several studies have looked at the content and characteristics of reviews. Sen and Lerman (2007) examine the impacts of review valence and suggest negative reviews garner more attention from consumers than positive ones. They also find that for hedonic products, readers are less likely to find negative reviews useful. Mudambi and Schuff (2010) find review valence and product type to be determinants of perceived review helpfulness. While these studies show that review valence and product type affect how consumers interact with reviews, the observations were collected at a fixed point in time, with no data on when consumers voted on a review. This limits the insights on the immediate effects of review characteristics and consumer interactions. The use of panel data by Ludwig et al. (2013) allows for an examination of the effects of affective review content and linguistic style matches over time on conversion rates. They also suggest that evaluating the order in which reviews appear on websites could be an opportunity for managers.

Dynamics of Consumer Rating Behavior

Examining the effect of previous ratings on future rating behavior provides a richer understanding of the influence of review and reviewer characteristics. For example, Schlosser (2005) finds consumer rating behavior is influenced by social pressures, and consumers alter their product evaluations based on other reviews they have seen. This demonstrates consumer reviewing behavior is susceptible to influences outside of the product itself, including information from other consumers. Godes and Mayzlin (2004) look at review timing as a factor that can influence reviewing behavior. They find dispersion (the breadth of discussion across online communities) of early reviews can influence the valence of future ratings, suggesting that review timing may be an important consideration. Other research evaluated the impacts of the review environment. Moe and Schweidel (2011) show that the valence of the rating environment can affect a consumer's decision on what to post, and whether to post at all. This demonstrates that the nature of review environment can determine whether or not a member of the community chooses to participate. Moe and Trusov (2011) examine the effects of previously posted ratings on future ratings, showing that the valence of a rating will be influenced by the valence of preceding ratings. Sridhar and Srinivasan (2012) find evidence of social influence in online review environments by demonstrating that the reviews of other consumers can ultimately impact how a reviewer rates a product by moderating the effects of product experience on review ratings. The literature clearly demonstrates social dynamics play a role in consumer rating behavior and motivates a further examination of the social dynamics involved in consumer interaction with reviews. Table 1 provides a summary of key literature on the streams of research concerning online reviews discussed here.

Table 1. Key Findings on Online Reviews

Study	Key Findings
Effects of Reviews on Sales and Strategy	
Chevalier and Mayzlin (2006)	Online book ratings affect consumer purchase behavior; impact of one star reviews greater than five star reviews.
Liu (2006)	Word of mouth can be used to explain movie box office revenue, but only in terms of volume, not valence.
Dellarocas, Zhang, and Awad (2007)	Adding online review metrics to sales forecasting models substantially increases forecasting accuracy.
Duan, Gu, and Whinston (2008)	Accounting for endogeneity, the valence of online reviews does not impact movie box office revenue, but volume does.
Zhu and Zhang (2010)	Product and consumer characteristics moderate online review influence on sales.
Tang, Fang, and Wang (2014)	The volume of neutral user generated content amplifies the effect of positive and negative user generated content on product sales.
Influence of Review Characteristics	
Sen and Lerman (2007)	Product type moderates the effect of review valence; consumers have a negativity bias for utilitarian products.
Mudambi and Schuff (2010)	Review extremity, review depth, and product type affect the perceived helpfulness of reviews.
Ghose and Ipeirotis (2011)	Review text can be used to understand consumer preferences for product features.
Schlosser (2011)	Two sided arguments (pros and cons) in online reviews are not always more helpful than one sided reviews; review extremity mediates this relationship.
Ludwig et al. (2013)	Affective content and linguistic style matches impact conversion rates.
Yin, Bond, and Zhang (2014)	There is a connection between the emotional content of a review and its perceived helpfulness; anxious reviews are considered to be more helpful than angry reviews.

Table 1. Key Findings on Online Reviews – Continued

Dynamics of Consumer Rating Behavior	
Schlosser (2005)	Online posters are influenced by the negative opinions of others in the environment and adjust their activity because of it.
Forman, Ghose, and Wiesenfeld (2008)	Disclosure of reviewer identity information influences the behavior or future reviewers.
Moe and Trusov (2011)	Social dynamics impact rating behavior and product sales.
Moe and Schweidel (2011)	Identified factors in the rating environment that influence incidence of posting behavior.
Sridhar and Srinivasan (2012)	Online reviews of other consumers moderate the effects of product experience, failure, and recovery on product ratings.
Zhao et al. (2013)	Consumers learn more from online book reviews by others than they do from their own experience with similar products.

Review Interactivity

Alba et al. (1997) conceptualize interactivity as “a continuous construct capturing the quality of two-way communication between two parties” (p. 38). Consumer to consumer interactivity has been identified as a specific type of interaction within online communities (Yadav and Varadarajan 2005). Consumer interactions with online reviews serve as a form of communication between the reader and reviewer through comments, likes, and votes. As the manner and name of these interactions vary from site to site, these types of visible interactions will be referred to as “consumer-review interactions” (CRI). The reasons consumers seek, use, and respond to information are important in understanding responses to communications (Stewart and Pavlou 2002). It has been shown that this type of non-transactional behavior can add value to a firm by encouraging purchases, referrals, and knowledge development (Kumar et al. 2010). To understand why consumers may be inclined to interact with online reviews posted by others, it is important to look at research that may influence this behavior. Moe and Trusov (2011) and Sridhar and Srinivasan (2012) look at social dynamics by examining elements of information posted in review forums that influence other consumers. While this is likely to be an important consideration, the norms present in the review environment are another aspect of social influence that needs to be considered.

Conceptual Development and Research Hypotheses

Community Norms

Virtual communities represent social networks that can affect consumer behavior (de Valck, van Bruggen, and Wierenga 2009). Kozinets (1999) suggests that virtual communities can refer to online groups of people who share norms and may actively enforce certain moral standards. Online communities where reviews are written and viewed have norms and standards regarding what is acceptable and appropriate in the forum. While not intended to be exhaustive, Table 2 provides a summary of the key literature concerning online communities.

Table 2. Key Findings on Online Communities

Study	Key Findings
Influence of Online Communities on Marketing Messages and Strategy	
Kozinets (1999)	Marketing in virtual communities must center on consumers being proactive and communally influenced.
Kozinets (2002)	Netnography is developed as a method to study online communities.
Mathwick (2002)	Investment in infrastructure allowing consumers to connect beyond transactional levels leads to future consumer loyalty.
Dwyer (2007)	Content that is highly valued by an online community attracts attention regardless of who created it.
de Valck, van Bruggen, and Wierenga (2009)	Virtual communities serve as reference group for consumers that differ from traditional reference groups and influence consumer decision making.
Kozinets et al. (2010)	Marketing messages are systematically altered by elements present in online communities.
The Influence of Online Community Characteristics, Features, and Participants	
Bagozzi and Dholakia (2002)	Positive emotions, social identity, and desires drive consumer participation in online communities.
Brown, Broderick, and Lee (2007)	Online communities can serve as a proxy for individual identification among its members.
Ren, Kraut, and Kiesler (2007)	The design and structure of online communities can lead to common identity or interpersonal bonds among community members.
Moon & Sproull (2008)	Feedback provided in online communities can lead to higher quality contributions and longer participant duration.
Seraj (2012)	Intellectual value, social value, and cultural value are features of online communities that create value for community members.
de Almeida et al. (2014)	Diversity on online communities facilitates learning, but hinders social identification.

Most customer review forums allow for free-form text reviews where the content can vary widely. Retailers that provide these forums set rules and offer best practices designed to guide review content. While severe violations such as obscenities and violence are actively monitored and removed from the website, reviews that more subtly violate these rules and guidelines generally remain on the site (Zhang, Craciun and Shin 2010). This content that violates norms but is not egregious enough to warrant removal may still face reaction from the community. The nature of these norms and how they are enforced are critical in understanding the dynamics of CRI within the community. The specific textual content of each review is evaluated against what is deemed acceptable in the online review.

Amazon.com for example, forms an online community where norms affect reviewer behavior (Forman, Ghose, and Wiesenfeld 2008). Ren, Kraut, and Kiesler (2007) use common identity theory to establish some behavioral outcomes typical of online communities. They put forth that the type of online community present is likely to influence the content of the discussion within the community. Common bond communities are based on attachments among group members while common identity communities are based on attachments to the identity of the group (Prentice, Miller, and Lightdale 1994).

Online review communities typically provide minimal personal information about their members and offer limited opportunities for one-to-one communication between their members, suggesting that members are more drawn to the identity of the community as a whole than individual bonds with other members. Many popular review forums, like Amazon and Yelp, are identity-based online communities that have established norms. Forman, Ghose, and Wiesenfeld (2008, p. 309) acknowledge the common identity nature of the Amazon community: “Amazon reviewers and members are cooperatively interdependent, with reviewers providing information that other members value, and members granting peer recognition that reviewers value. Interdependence and a common purpose facilitate common identity rather than common bond". This identity based attraction is reinforced by recognition, acknowledgements, and incentives provided by the community organizers. Popular review sites like Amazon, Trip Advisor, and Yelp all acknowledge top reviewers by providing "badges" or displaying reviewer rankings which provides confirmation and recognition of group membership. These rewards reinforce a reviewers membership in the Amazon review community and are a symbol that members can take pride in. Members of this type of community depend on others to provide them with the CRI they need to retain their status (Forman, Ghose, and Wiesenfeld 2008; Ren, Kraut, and Kiesler (2007). This provides motivation for some reviewers to produce highly interactive reviews.

Off-topic discussion is identified as an issue that is undesirable in identity based communities, and many online communities are designed with provisions to keep content on topic (Ren, Kraut, and Kiesler 2007). Amazon has rules for posting reviews, explicitly stating that off-topic information is not allowed in their review forum. Non-retail review communities like Yelp.com also provide guidelines to keep review content on topic and provide the following instruction for members: "Please make sure your contributions are relevant and appropriate to the forum. For example, reviews aren't the place for rants about a business's employment practices, political ideologies, extraordinary circumstances, or other matters that don't address the core of the customer experience" (Yelp.com).

Governing through social networking reinforces behavioral expectations in brand communities (Schau, Muniz, and Arnould 2009). In an online context, communal norms govern relationships between community members (Mathwick 2002) where contributors communicate the importance of group norms to each other (Moon and Sproull 2008). Kozinets et al. (2010) show that over time, conventional practices used in online communities form the basis of implicit social norms, and there can be a community response from violations of those norms. Since personal contact information is not available in most review forums, interacting with a review itself is often the only recourse a community has to share their concerns about norm violations. By voting or commenting on a review, members of the community share these concerns about content with the reviewer and the larger community. Since CRI are the most practical form of community response to online reviews that violate community norms, it is hypothesized that:

H1: Violations of established review community norms within the content of a review are positively associated with the level of consumer-review interactions (CRI).

While a violation of community norms in itself is likely to draw a reaction from the community through CRI, who is writing the review is likely to moderate this effect. Prior research has shown that aspects of a reviewer can influence reactions to a review they write. Forman, Ghose, and Wiesenfeld (2008) demonstrate that information disclosed by a reviewer can impact the perceived helpfulness of their reviews. Naylor, Lamberton, and Norton (2011) further argue that identity disclosure may play a role in the persuasiveness of a review dependent upon perceived similarities between the consumer and reviewer.

Another reviewer characteristic that can influence CRI is the reviewer's ability to write reviews that garner attention. Some reviewers are better and more experienced at writing engaging reviews that draw attention from the community. When a skilled reviewer produces a review that violates a community norm, the attention it receives will be amplified because they are skilled at drawing a response from the community. While norm violations themselves are likely to result in more CRI, this effect is likely to be exaggerated when the norm-violating review is written by a skilled reviewer. Because reviewers with a history of high interaction rates may be better able and more motivated to garner attention, norm violating reviews written by these authors will have amplified levels of CRI. Therefore it is hypothesized that:

H1a: Reviews that violate community norms will receive more (less) consumer-review interactions (CRI) when written by a more (less) skilled reviewer.

Visual Prominence

While community dynamics are likely to be part of the story, a consumer's opportunity to simply view an online review is an important aspect to consider when examining drivers of CRI. The physical context of where a review is located on a website may help drive the community's reaction to it. Some products draw hundreds or even thousands of reviews by customers, and in this crowded atmosphere, all reviews are not given an equal opportunity to be viewed. At any given time, the prominence of a review and how consumers respond to its prominence will drive CRI. Online reviews on major retail websites are usually arranged as an ordered list, similar to what is seen with web search results. For example, Amazon places the most helpful reviews at the top of the page, making those reviews the most visually prominent. Other popular review sites like Tripadvisor.com place ten reviews on the first page of results. This similarity to web search results suggests that online reviews that appear in more prominent page positions are more likely to be used than other reviews.

Research concerning consumer Internet search behavior and online advertising helps to explain the impact of review location within a retailer's website on the level CRI generated. A significant portion of users examine only the first page of search results (Spink et al. 2001); that number can be as high as 70 percent (Jansen and Spink 2006). When considering shoppers, the number of consumers who only select offers from the first screen can drop to nice percent (Brynjolfsson, Dick, and Smith 2010). The rank of results drives user propensity to click on the pages; higher ranked and positioned results receive more clicks than lower ranked results (Pan et al. 2007). Eye tracking studies of shopping sites have shown that consumers pay less attention to subsequent pages relative to the first one viewed (Pan et al. 2004). Part of the tendency to focus on the easily accessible first page of results is likely due to the high cost consumers associate with making just a few clicks (Agarwal, Hosanagar, and Smith 2011; Hann and Terwiesch 2003). The probability of an item being clicked depends on its relevance and its position on the page (Craswell et al. 2008). These studies suggest that the more visually prominent information on the web is, the more attention it will get with consumers. Therefore, it is hypothesized:

H2: The visual prominence of a review is positively associated with the level of consumer-review interactions (CRI)

As with norm violations, reviewer skill is likely to moderate the impact visual prominence has on CRI. On sites like Amazon, the visual prominence of reviews is in part determined by the level of interactivity a review receives. Reviews that receive a high volume of helpful votes are more likely to appear on the product or review landing page, increasing their visual prominence. Amazon in turn rewards members of the community for producing reviews that generate high amounts of positive interactions. Badges such as “Top 500 Reviewer” are awarded to reviewers based on their ability to generate helpful and engaging reviews. While some reviews that appear in visually prominence positions may be written by reviewers who are skilled at generating CRI, others may appear in prominent positions as a product of timing or chance. Among the visually prominent reviews, those written by reviewers who have a history of generating high levels of CRI are likely to have this attention amplified, resulting in more CRI. Therefore it is hypothesized that:

H2a: Reviews that are visually prominent will receive higher (lower) levels of consumer-reviewer interactions (CRI) when written by a more (less) skilled reviewer.

Social Dynamics

Moe and Trusov (2011) establish that social dynamics can influence a reviewer's evaluation of a product in the rating environment. It is also likely that similar social dynamics can impact how shoppers evaluate a review that has already been posted. Using information observed from others to formulate a decision may be an important driver of CRI. When a consumer sees that a review on Amazon has been considered helpful by 120 out of 125 voters, the consumer may accept that the review is helpful without expending additional cognitive effort to determine her own opinion about the review. Seeing a large number of total votes relative to other reviews has informational value, and this informational influence may encourage other consumers to cast a vote on the review too.

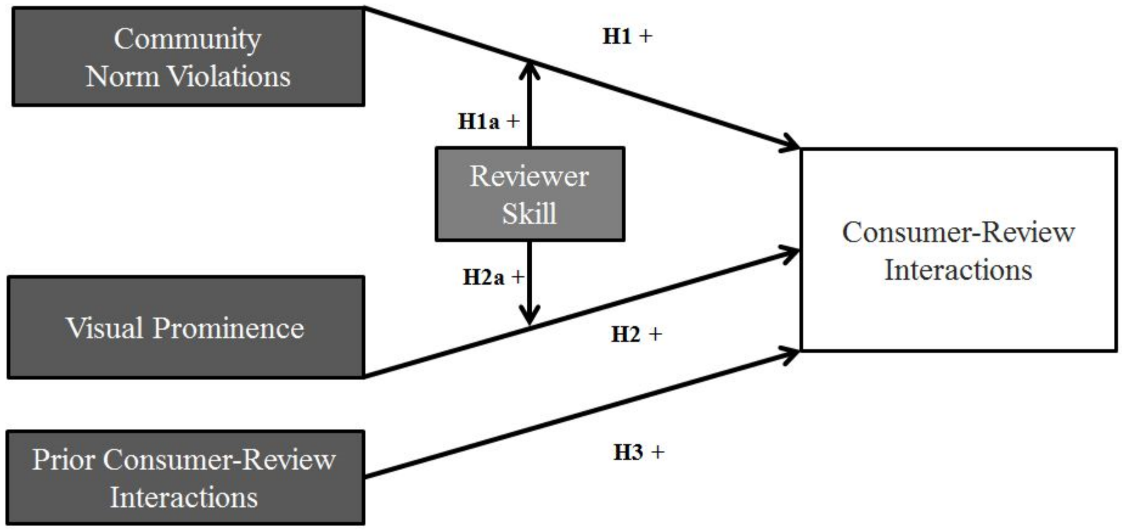
An information cascade occurs when individuals imitate or follow the behaviors of others rather than utilize and process their own information (Huang and Chen 2006). When this occurs, barring interruption, future individuals will continue to ignore other information, and will instead mimic the behavior of those before them (Bikhchandani, Hirshleifer, and Welch 1992). Informational cascades can start with a very small amount of information and be used to explain the rapid spread of a behavior (Bikhchandani, Hirshleifer, and Welch 1992). Also known as herding behavior, this has been shown to influence online consumer product choice (Chen 2008; Huang and Chen 2006). Sundar and Nass (2001) describe a "bandwagon" heuristic that was a key driver of ratings provided by online news readers, suggesting: "when available, information about how others react to a message is used by individuals judging the same message" (p. 68).

This type of behavior will have a similar effect on CRI. If a handful of consumers vote or comment on a particular review in greater proportion than others, an informational cascade could begin and create momentum for further interactions with that review. Liu et al. (2007) explain this "rich-get-richer" effect as a result of time and opportunity for reviews to be seen. By observing reviews at regular intervals over a period of time, it could be shown how this dynamic effect builds and impacts future CRI. Information cascades and herding behavior suggest that prior customer interactions with reviews will serve as a cue to consumers promoting additional CRI. Thus, it is hypothesized that:

H3: The amount of consumer-review interactions (CRI) in prior periods is positively related to the amount of consumer-review interactions (CRI) a review receives in a given period.

To summarize, CRI are affected by several factors concerning review content and context working simultaneously within the review environment. Violations of community norms, visual prominence, and social dynamics all work together to help explain the dynamics of online review interactivity. Refer to Figure 1 for an illustration of the conceptual model.

Figure 1: **A Conceptual Model of Consumer-Review Interactions**



Methods and Data

A major goal of this research is to evaluate drivers of CRI as occurring in real time within an established online community. To accomplish this, the platform, timing, and product category all need to be carefully considered. The Amazon.com reviewing platform was selected due to its scale and because it has been a successful source of data for a number of important studies concerning online reviews (Chevalier and Mayzlin 2006; Forman, Ghose, and Wiesenfeld 2008; Mudambi and Schuff 2010).

To best study this dynamic environment in real time, it is important to start capturing data as soon as the product is released into the environment. Using archival data or observing reviews after a product has been on the market for a period of time can miss out on important time-dependent information. For example, a review may be posted on the first day a book goes on sale. On that first day, the review in question may appear in a highly prominent position on the retailer's website. By the second day however, that same review may be located in a completely different location on the site, altering the chances a shopper may encounter it, and therefore interact with it. Due to this continual change in the review environment, collecting data at a static point would compromise the ability to observe these critical changes in the environment that may influence how a consumer views and potentially interacts with a review. To account for these changes in placement over time, reviews data can be collected continually from the time they are posted. This reduces the possibility of missing important environmental factors that could affect CRI.

To collect review data at the time of product release, products need to be identified prior to release. Books were chosen as the product for this study because book release dates are often announced far in advance. Also, books have been used successfully in several important studies involving online reviews (Berger, Sorensen, and Rasmussen 2010; Chevalier and Mayzlin 2006; Forman, Ghose, and Wiesenfeld 2008).

Data Collection

A sample of 27 books was selected for observation before their release. The sample is a representative cross section of subject, genre, and price. Fiction and nonfiction books were included, with categories ranging from romance, to children's literature, to public finance. Release prices ranged from $7.99 to $20.58. The first 30 reviews posted for each book were included in the study. This number ensured both the feasibility of data collection and a diverse sample of reviews for each book. For each of the first 30 reviews posted for each book, detailed data were collected for 28 days following the release of the book. Data on the website environment, the books, the reviews, and the reviewers were collected every 24 hours for the first 14 days after release, and every 48 hours for the next 14 days. A pilot study indicated that site activity was highest during the initial 14 day period following a book release, then slowed substantially, motivating the less frequent data collection during the second 14 day period.

The data collection resulted in 22 observation periods over 28 days for 30 reviews of each of the 27 books, for a total of more than 11,000 observations. This data collection method enables an analysis of how CRI change over time as variables in the environment change. This would not be possible by collecting static, point in time, cross-sectional data.

Variables

The dependent variable selected for the analysis is the number of customer-review interactions (CRI) during a given observation period. Consumers can interact with a review on Amazon by voting or commenting on it. While these actions are distinct, they both represent an observable interaction with a review and represent CRI in this context. CRI are calculated by adding the number of votes and comments received by a review during an observation period. Defining the dependent variable in this manner ensures that the impact of dynamic variables in a review's environment is properly matched with the corresponding CRI of the same period. CRI could also be thought of as the amount of incremental interaction a review receives each period. The independent and control variables included in the study were also collected in each of the 22 observation periods. The explanatory variables and their definitions are provided in Table 3.

Table 3. Chapter 2 Variable Descriptions

Variable	Description
	Dependent Variable
Consumer-Review Interactions (CRI)	The combined amount of votes and comments received by a review in a given observation period.
	Independent Variables
Visual Prominence	Whether or not the review appeared on the product landing page or customer review landing page in a given observation period.
Norm Violation	Whether or not the review content violates community norms as defined in Appendix 1.
Prior Consumer-Reviewer Interactions (Prior CRI)	The cumulative number of votes and comments received by the review prior to the given observation period.
Reviewer Vote Rate	The reviewer's historical rate of receiving votes on reviews. Calculated by examining all product reviews posted by the reviewer prior to the data collection for this study. Computed as the total votes received on prior product reviews, divided by the number of product reviews posted.
	Control Variables
Star Rating	The star rating of the review.
Net Emotion	A measure of the valence of the written content contained in a review. Calculated as the positive emotion word count minus from the negative emotion word count and divided by the total word count of each review.
Helpful Score	A measure of review helpfulness. Computed to emulate how Amazon ranks helpfulness. Computed as (Helpful Votes - 2 * Not Helpful Votes).
Observation Period	Whether or not the observation was recorded in the first five periods after the product was launched.
Sales Rank	The Amazon sales rank in the book category in the given observation period. Inversely related to sales; the greater the sales the lower the sales rank (bestselling book would have sales rank of 1).

Static and dynamic characteristics of each of the first 30 reviews were captured during each observation period. The first variable, Visual Prominence, describes where on the website the review is positioned in a given observation period. Amazon places certain reviews in special positions to make them more visually accessible to consumers. The reviews selected to be in these critical positions continually change, further highlighting the importance of continuous data collection. There are two key pages that need to be considered when analyzing visual prominence: the product landing page and the review landing page. The product landing page is the first page consumers see when they click on and view the product. The review landing page is the first page consumers see when they decide to view the reviews available for a product. The default display of the reviews on both landing pages present the ten most helpful reviews as determined by Amazon. As most searchers view only the first page of results (Jansen and Spink 2006), whether or not a review appeared on either the product landing page or the review landing page in a given observation period defines the variable Visual Prominence. Visual Prominence was coded dichotomously (Visually Prominent = 1, Not Visually Prominent = 0). Since reviews placed in in these highlighted positions continually change in response to CRI, this emphasizes the importance of recording these positions daily.

Amazon provides guidelines for what is not allowed to be posted in a review and removes extreme violations. More subtle violations break the norms of the online community and elicit a response from the community in the form of review interactions. The Norm Violation variable was determined by an evaluation of the review content by two research assistants. The evaluators were provided with the text of each review and a set of instructions (see Appendix 1). They were asked to independently read the text of each review and determine if the review violated the community norms as described in the instructions (yes/no). In the event of a disagreement between the evaluators, a third evaluator was used to break the tie. The Norm Violation variable indicates whether or not the review content has violated what is deemed acceptable by Amazon and the community. The variable was coded 1 for a violation and 0 for no violation. It is expected that Norm Violations will be positively associated with CRI.

To capture the propensity of individual reviewers to write reviews that garner a high level of CRIs, historical information about the author of each review considered in the analysis was collected. Amazon provides historical reviewing activity for every reviewer. The rate at which reviewers historically generate CRI serves as a proxy for reviewer skill, represented by the variable Reviewer Vote Rate. Reviewer Vote Rate is determined by examining a reviewer's history and is the average number of votes received per review written by the reviewer prior to posting the review being observed in the analysis. Reviewer Vote rate was included to account for the history and potential skill of a reviewer to write highly interactive reviews.

When viewing a particular review, shoppers are also able to see an indication of previous interaction with a review. At the top of each review the cumulative number of votes received is displayed (i.e. 12 out of 15 people found the following review helpful). At the bottom of each review the cumulative number of comments the review has received is displayed. Combined, these two figures show the amount of prior interactions shoppers have had with the review and make up the variable labeled as "Prior CRI". For example, when "12 out of 15 people found the following review helpful" and the review had two comments, Prior CRI has a value of 17. Like Visual Prominence, Prior CRI varies during the life of a review, further highlighting the need for frequent observation.

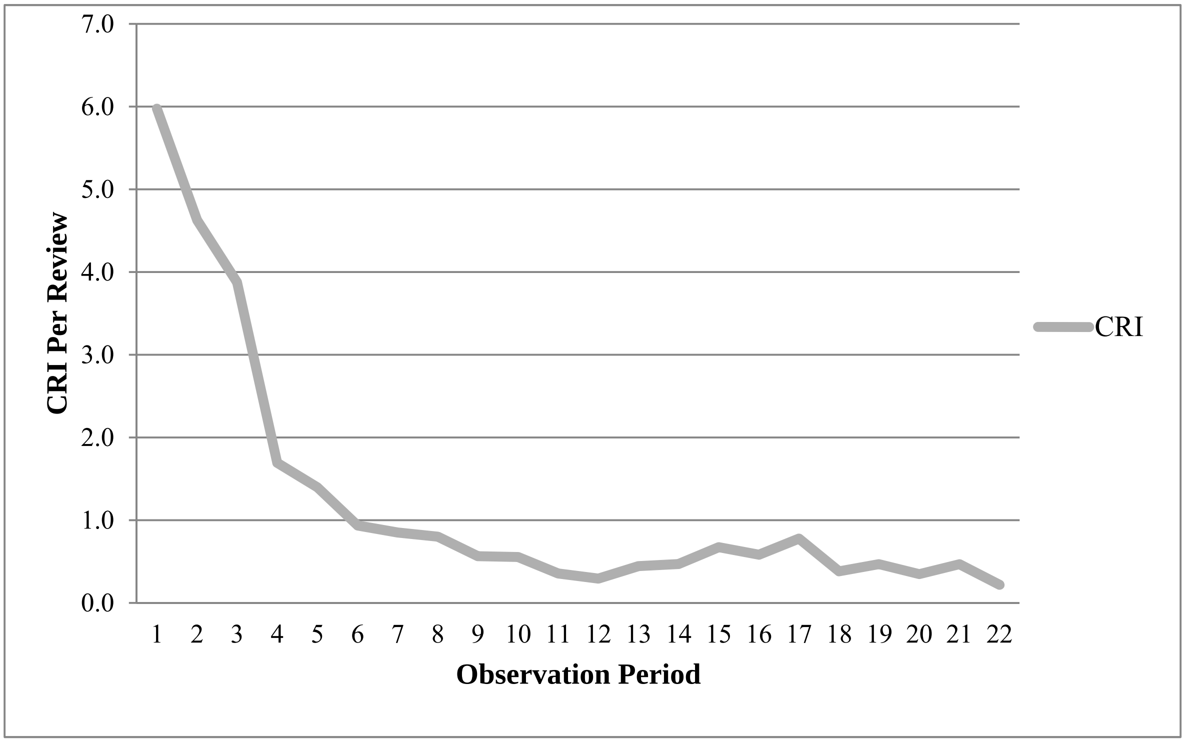
Several control variables were collected to account for other factors that could potentially influence the level of CRI. Two variables were used to address the valence of each review. The star rating assigned by the reviewer (one to five stars) is commonly used as a measure of valence in research on online reviews (see Mudambi and Schuff 2010; Sen and Lerman 2007) and has been included in this analysis, represented by the variable Star Rating. While star rating provides a direct signal of a reviews' valence, it is possible that the relative negativity contained in the text of the review does not perfectly align with the star rating assigned to it (Rill et al. 2013). To address this issue, an analysis of the text of each review was completed to derive a measure of the valence of the written content.

Reading text containing affective content can influence consumer behaviors, suggesting that the level of affective content present in a review could influence a consumer's reaction to it (Lau-Gek and Meyers-Levy 2009; Ludwig et al. 2013). To determine the intensity of affective content contained in the text of each review, a procedure similar to that of Ludwig et al. (2013) was used. The textual content of each review was analyzed using Linguistic Inquiry and Word Count (LIWC), a program originally designed to analyze emotional writing (Pennebaker et al. 2007). The dictionaries contained in LIWC offer reliable estimations of the ratings they extract and their use has been validated in studies concerning online review content (Ludwig et al. 2013; Pennebaker et al. 2007). Two LIWC dictionaries (positive emotion and negative emotion) were used to identify the proportions of affective content contained in the text of each review. The program dictionaries identify the number of words associated with positive emotion (e.g., love, nice, sweet) and negative emotion (e.g., hurt, ugly, nasty) contained in each review and reports them as a percentage of the total word count for that review (Pennebaker et al. 2007). To make a single measure representing the total directionality of the affective content present in each review, the negative emotion word count was subtracted from the positive emotion word count and divided by the total word count of each review. This created a summary score for each review which falls between -1 and 1. For example, a 100 word review which contains 10 positive emotion words and five negative emotion words would have a score of .05 (i.e. 10/100 – 5/100 = .05). This measure of net emotion, referred to as the Net Emotion variable, serves as a proxy for the valence of the written text of each review in the analysis. It is expected that both measures of valence will be inversely related to CRI (negative reviews will have more CRI than positive reviews).

The Helpful Score is derived in a way that closely mirrors the method used by Amazon to rank review helpfulness, as evidenced by position on the page. Helpful Score is determined by the following calculation: (Helpful Votes – 2*Not Helpful Votes). Using this method provides an advantage over using a count of helpful votes or the percentage of helpful votes as used in some studies involving review helpfulness (see Forman, Ghose, and Wiesenfeld 2008; Ghose and Ipeirotis 2011; Mudambi and Schuff 2010). The Helpful Score as calculated here mimics the method used by Amazon to evaluate review helpfulness meaning that this study uses helpfulness in the same manner it is judged in the review environment. It is expected that Helpful Score is positively related to CRI.

There are two main reasons why it is important to note at what point during the data collection a particular observation was made. The first is a simple competition effect. The longer books are on the market, the greater the opportunity for consumers to post reviews. Having more reviews available to choose from will likely lower the probability that a consumer will choose to interact with one review in particular. Second, the rate of CRI is not level throughout all periods. This point is illustrated by examining the average number of CRI per review in each observation period. Figure 2 illustrates that the average CRI rate is substantially higher in the periods immediately after a books release, due to the higher level of aggregate CRI and to the lower amount of total reviews that exist in the initial periods. After the fifth observation period, the average CRI per review slips below one and remains relatively flat for the duration of the data collection period. This suggests that time may have a significant influence on generating CRI. To control for time, a dichotomous variable, Observation Period, was included in the analysis. Observations that took place in the first five periods after release were coded 1, and those in periods six and after were coded 0. It is expected that Observation Period is positively related to CRI.

Figure 2. Average Customer-Review Interaction (CRI) per Review by Period



Sales Rank was included in the analysis to account for the popularity of the book. If a book is a best seller, it is likely to have more shoppers viewing it, thus providing more opportunities for CRI. To control for this effect, the Amazon sales rank for each product was recorded in each observation period. Numerically, sales rank is inversely related to popularity. For example, a book with a sales rank of five is the fifth most popular book in a given category and would be more popular than a book with a sales rank of 50. As such, it is expected that Sales Rank will be inversely related to CRI.

Analysis

The data collection yielded 11,209 observations, comprised of 727 reviews across 27 books. Small amounts (< 1%) of missing data occurred due to instances where sales rank or author review history was unavailable. These observations were excluded from the analysis leaving 11,058 observations in the dataset. The descriptive statistics for each variable are shown in Table 4. The CRI variable mean was .78. This means that, on average, each review had less than one CRI per period. The range (min: 0, max: 120) and standard deviation (2.98) indicate that the level of CRI is indeed volatile and further suggests certain reviews may have a tendency to garner more attention. The mean of Visual Prominence indicates that in 43% of the observations collected, the review was located in a visually prominent position. The mean of Norm Violation indicates that 10% of the reviews included in the analysis violated the norms of the community. Reviewer Vote Rate has a mean of 3.56 indicating that on average, the authors of the reviews observed in the study had received 3.56 votes for each of their reviews written prior to this study. Consistent with prior studies the mean star rating (4.20) suggests that reviews generally have a positive valence. The positive mean of the Net Emotion variable (.036) also suggests a generally positive valence of review text.

Table 4. Chapter 2 Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
		Dependent Variable		
Consumer-Review Interactions (CRI)	.78	2.98	0	120
		Independent Variables		
Visual Prominence	.43	.50	0	1
Norm Violation	.10	.30	0	1
Prior CRI	9.56	23.50	0	411
Reviewer Vote Rate	3.56	7.54	0	194
		Control Variables		
Star Rating	4.20	1.22	1	5
Net Emotion	.036	.043	-.154	.273
Helpful Score	-1.37	22.40	-209	286
Observation Period	.11	.32	0	1
Sales Rank	737.23	1474.20	9	19470

N = 11,058

Two mixed linear models including a random effect for Product were constructed to test the hypotheses presented. Given that the reviews are observed from a group of 27 books it is likely that there is correlation among observations drawn from the same book. This presents an endogeneity issue in that the individual observations are not independent due to being drawn from the same subject (book). Introducing the Product variable as a random effect in the model helps to address this issue as the products selected for analysis are a small sample of a large population of books available for sale on Amazon.

As noted earlier, the dependent variable, CRI, represents the total number of votes and comments a review received since the prior observation period. More specifically, CRI^{br}_{ijt} is defined as the number of votes and comments received by review i of book j at time t. Since the data was collected at frequent intervals rather than truly continuously, it is unclear at exactly what point in the observation interval a CRI may have occurred. For example, when considering a particular review, an observed CRI could have occurred immediately after the prior observation period, immediately before the current observation period, or somewhere in between. To address this issue, models using both lagged and non-lagged independent variables were constructed. Model 1 uses independent variables collected in the same observation period as the dependent variable, or time t. The data set for testing Model 1 contained 11,058 observations and is specified as:

$$\text{CRI}^{br}_{ijt} = \beta_0 + \beta_1\text{Visual Prominence}^{br}_{ijt} + \beta_2\text{Norm Violation}^{br}_{ij} + \beta_3\text{Prior CRI}^{br}_{ijt} + \beta_4\text{Reviewer Vote Rate}^{br}_{ij} + \beta_5\text{Reviewer Vote Rate}^{br}_{ij} * \text{Visual Prominence}^{br}_{ijt} + \beta_6\text{Reviewer Vote Rate}^{br}_{ij} * \text{Norm Violation}^{br}_{ij} + \beta_7\text{Star Rating}^{br}_{ij} + \beta_8\text{Net Emo}^{br}_{ij} + \beta_9\text{Helpful Score}^{br}_{ijt} + \beta_{10}\text{Observation Period}^{br}_{ijt} + \beta_{11}\text{Sales Rank}^{b}_{it} + \gamma\text{Product} + \varepsilon$$

Model 2 uses independent variables that were present at the prior observation period, or time t-1. Only the independent variables that can change over time were altered in Model 2. By including both models, concerns about potential changes in the environment that may have occurred between observation periods are alleviated. The dataset was adjusted for use in Model 2 by removing observations that took place the first time a review was available, and by removing reviews that had only one observation as these data points have no information from prior observations. The remaining data set for Model 2 contained 10,337 observations and is specified as:

$$CRI^{br}_{ijt} = \beta_0 + \beta_1 \text{Visual Prominence}^{br}_{ijt-1} + \beta_2 \text{Norm Violation}^{br}_{ij} + \beta_3 \text{Prior CRI}^{br}_{ijt-1} + \beta_4 \text{Reviewer Vote Rate}^{br}_{ij} + \beta_5 \text{Reviewer Vote Rate}^{br}_{ij} * \text{Visual Prominence}^{br}_{ijt-1} + \beta_6 \text{Reviewer Vote Rate}^{br}_{ij} * \text{Norm Violation}^{br}_{ij} + \beta_7 \text{Star Rating}^{br}_{ij} + \beta_8 \text{Net Emo}^{br}_{ij} + \beta_9 \text{Helpful Score}^{br}_{ijt-1} + \beta_{10} \text{Observation Period}^{br}_{ijt-1} + \beta_{11} \text{Sales Rank}^{b}_{it-1} + \gamma \text{Product} + \varepsilon$$

Results

The results of the regression analysis for both models are shown in Table 5. Reviews that violate community norms were shown to be significantly positively associated with CRI in both models as indicated by the Norm Violation variable ($p<.001$). This provides strong support for Hypothesis 1.

The interaction term, Norm Violation*Reviewer Vote Rate was positive and significant ($p<.0001$). This suggests that when a norm violation is committed by a skilled reviewer, CRI will be higher than if the norm violation was committed by a less skilled reviewer. Model 1 predicts that (holding all else constant) at the mean level of Reviewer Vote Rate, a review that violates norms would expect to receive 2.24 CRI in a given period. At one standard deviation above the mean of Reviewer Vote Rate, a review that violates norms would expect to receive 2.74 CRI in a given period, a 22.3% increase. At a Reviewer Vote Rate of zero (a negative vote rate is not possible), the model predicts 1.79 CRI in a given period, a 20% decrease. When looking at a review that does not violate norms, the predicted CRI increase at one standard deviation above the Reviewer Vote Rate mean is only 1% while the expected decrease at a Reviewer Vote Rate of zero is only 2%. This illustrates the predicted interaction effect and provides strong support for Hypothesis 1a.

The Visual Prominence variable was positive and highly significant ($p<.0001$) in both models, providing strong support for Hypotheses 2. The interaction term, Visual Prominence*Reviewer Vote Rate was also positive and significant ($p<.0001$) in both models suggesting that visually prominent reviews receive more CRI when they are written by a skilled reviewer. This provides support for Hypothesis 2a.

Surprisingly, Prior CRI was not significant in Model 1, but was positive and significant ($p<.0001$) in Model 2. The mixed result is surprising, but is intuitive upon further inspection. The dataset used in Model 1 included all observations, including those where there was only a single observation and those where it was the first time a review was observed. In both cases these observations have no opportunity for prior CRI, which eliminates the possibility of consumers being influenced by past behavior. The dataset for Model 2 excludes these observations from the analysis and only includes reviews that had a prior opportunity for CRI. Model 2 offers a better specification to test Hypothesis three by excluding review observations that had no possibility of prior viewing. The positive and significant coefficient of Prior CRI in Model 2 suggests that in cases where there has been prior opportunity for consumers to view a review, Prior CRI are positively associated with CRI in the current period, providing support for Hypothesis 3.

The control variables generally performed as expected. Star Rating was found to be negative and significant ($p<.0001$) in both models, suggesting that review valence is inversely related to CRI. This result is consistent with previous research concerning review valence in that it suggests reviews with lower star ratings receive more attention from consumers. Net Emotion, measuring the valance of written text of the reviews, offered mixed results. In Model 1, Net Emotion was negative and significant ($p<.05$), further suggesting consumer preference for negative reviews. In Model 2, however, Net Emotion is insignificant. This mixed result suggests that evaluation of different measures of review valence and their influence may be an interesting area for future research. Helpful Score was shown to be positive and significant ($p<.0001$) in both models indicating that more helpful reviews draw more CRI. Observation Period was also positive and significant ($p<.0001$) in both models, suggesting that reviews receive more CRI during the first five days following product launch.

Table 5. Regression Results – Dependent Variable: Consumer-Review Interactions (CRI) [a]

	Model 1			Model 2		
Variable	Coefficient	Std Error	Sig	Coefficient	Std Error	Sig
Intercept	1.40	.065	<.0001	1.01	.061	<.0001
Visual Prominence	.068	.007	<.0001	.043	.007	<.0001
Norm Violation	.045	.010	<.0001	.033	.010	.0006
Prior CRI[a]	-.003	-.003	.5345	.062	.005	<.0001
Review Vote Rate[a]	.096	.007	<.0001	.079	.007	<.0001
Norm Violation*Reviewer Vote Rate	.040	.007	<.0001	.039	.006	<.0001
Visual Prominence*Reviewer Vote Rate	.040	.005	<.0001	.033	.005	<.0001
Star Rating	-.091	.004	<.0001	-.079	.004	<.0001
Net Emotion	-.271	.119	.0225	-.081	.115	.4810
Helpful Score	.002	.000	<.0001	.003	.000	<.0001
Observation Period	.164	.008	<.0001	.101	.008	<.0001
Sales Rank[a]	-.105	.008	<.0001	-.079	-.079	<.0001
N	11,058			10,337		
R-squared	.383			.364		
Adj R-squared	.383			.364		
-2 Log Likelihood	14,267			11,857		
AICc	14,295			11,885		
BIC	14,397			11,986		

[a]Indicates variables were log+1 transformed

Discussion

Discussion and Implications

This research builds on the online review literature in several ways. By looking at customer-review interactions in the context of the dynamic rating environment, this research fills a gap and takes a step toward understanding what generates activity in an online review community. Of particular interest is the introduction of community norms as a factor that influences consumers in the review environment. The results presented here show that the dynamics behind review interactivity range from simple to complex. Findings also highlight that there is likely no single factor behind CRI. A system of factors that includes norms and visual prominence, works dynamically to influence how consumers interact with reviews.

The most significant theoretical contributions of the research work are in the area of social dynamics and online content, and in the introduction of community norms as a factor in online reviews. These results related to social dynamics build upon the work of Moe and Trusov (2011), Moe and Schweidel (2011), and Sridhar and Srinivasan (2012) by demonstrating that the review environment impacts consumer response to reviews in multiple ways. Although previous research (see Kozinets 2010, Forman, Ghose, and Wiesenfeld 2008) has examined norms in online communities, the influence of norms on review interactivity has not been examined in past conceptual frameworks. The results from this study show that norm violations are associated with a community response in the form of customer-review interactions. Introducing norm violations as an influential factor in communication levels within online review communities adds to the conversation and provides another dynamic to be considered in online review research.

Sometimes the simplest answers can be the most impactful. This seems to be the case when examining visual prominence of reviews. If shoppers are more likely to see a review, they are more likely to interact with it. While simple, this finding does help link research on internet search to online reviews, showing that consumers respond similarly to results in both formats.

While these findings add interesting contributions to the online review literature, the inclusion of the interactions with reviewer skill adds depth to the findings. Taking characteristics of the reviewer into account demonstrates that characteristics of the review and the review environment do not show the whole picture. Relatively less research has been done to examine characteristics of the consumers who write reviews and the impact they have on consumer behavior in online communities. This creates an area rich for future research.

The nature of the data collection used in this study also presents opportunities for future research. It is clear that the period shortly after a product is released generates more CRI than in periods later in the product's life. While this research demonstrates that these early periods to generate more interaction, it is not entirely clear why. Further research is need to examine what dynamics occur shortly after a product's release that lead to this phenomenon. The increased activity level upon product release could have interesting consequences for retailers.

This research fills a gap by showing that social dynamics, specifically the responses to violations of the review community's norms, can influence how consumers respond to previously posted reviews. This study examines these dynamics in real time, as the reactions unfold, and this manner is not frequently used in prior research. The longitudinal observation of reviews and the associated community responses in a changing environment offers a new way to study consumer generated online content. By examining a review, its characteristics, its environment, and the community response as it happens, this research adds insights concerning consumer behavior in online review communities.

The factors associated with consumer-review interactions have interesting potential managerial considerations as well. Reviews that are placed in visually prominent positions on retailers' websites receive more CRI. While it seems retailers like Amazon understand and value this, and place reviews that are the most helpful to their shoppers in these positions, the firms selling their products via Amazon may not see this advantageous. For example, book authors and publishers are likely to prefer not to have a negative review featured in a visually prominent position. When a negative review is deemed to be very helpful by the community, it will likely be pinned in a prominent position on the product page and/or review landing page creating a "digital book jacket" that is not as positive as the physical book jacket. As it has been shown here, the continued CRI these reviews are likely to receive indicate the attention paid to these reviews. Firms selling products via Amazon need to be prepared for the possibility of negative reviews becoming highly visible in the location their product is sold.

Violations of community norms were shown to be strongly associated with consumer-review interaction in this study. Retailers such as Amazon.com and review sites like Yelp.com should take note of this phenomenon as the implications are twofold.

On one hand, it indicates the existence of a strong review community that reacts and defends itself when reviews go outside what is generally deemed acceptable. On the other hand, it could indicate that the failure to more actively monitor reviews is enabling the creation of content that is not appreciated by active members of the community and could lead to dissatisfaction with the forum. Retailers would be wise to monitor the regulation of norms within their review communities.

Limitations

Further opportunities exist to enhance the data collection method established in this study. Using books exclusively calls into question possible effects product type may have on review interactivity. Mudambi and Schuff (2010) showed that the type of product (search vs. experience) moderated the effect of review depth on helpfulness while Sen and Lerman (2007) showed that product type (hedonic vs. utilitarian) moderated the negativity effect. Additionally, Huang, Lurie and Mitra (2009) find that product type (search vs. experience) influences the depth and breadth of consumer search when shopping online. It is possible that product type could be a factor that explains consumer-review interaction, but there is no clear indication of the nature of the effect. If a method were determined to identify the release dates of products in electronics and other product categories in advance, future research could examine this issue.

Beyond product type, there are other elements in the review environment that could be examined at a more granular level. Several variables concerning the author of a book could be examined to determine any possible relationship with interactivity. While this study includes the sales rank of a book as a measure to control for the relative popularity of a book, there may be aspects of the author that tell a more complete story. It is possible that certain authors may write controversial or otherwise stimulating books that inherently draw more interaction. A detailed analysis of historical reviewing activity of an authors' past titles could reveal such a trend. Similarly, some authors write books across multiple genres, which could provide a base of comparison for differences in review interactivity. First time authors or those without any review history could also be an interesting group to contrast as these authors have no expectations or prior commentary on their work, which could potentially lead to unique reviewing behavior and interactivity.

There are also additional aspects of the reviewers that could be considered in future research. On Amazon for example, there are reviewers who have written thousands of reviews across a wide range of product types. There are also conglomerates or groups that write reviews under one Amazon screen name despite the reviews actually being authored by multiple individuals. These types of reviewers may product content that contrasts with the casual user who has only written one or two reviews in their history. While this research introduces reviewer skill as a measure of a reviewer's history of writing interactive reviews, there may be differences in motivation and content between different types of reviewers. Future research could evaluate these differences and also look at how reviewers may alter their style and associated interactivity levels across product types.

The prominence of mobile technology and the ease at with consumers can take pictures could have consequences for online reviews and interactions with them. Many review sites now enable consumers to include images in their reviews. It is likely that reviews that include images may appeal to consumers differently than reviews with text alone which could impact interactivity levels. Chen and Ho (2015) for example find that online reviews that contain more images are perceived to be more useful to consumers. The inclusion of images in reviews, the volume of them, and the content of them, have the potential to influence review interactivity as they provide a richer level of content than text only reviews. Given the prominence of consumer picture taking, this may be an area ripe for additional research.

Since all of the observations collected were from one retailer, Amazon.com, it is possible that consumer interaction patterns could vary with other retailers. It is also possible that shoppers exhibit different behavior depending on the shopping season or time of year. Also, while it is likely that making observations every 24 to 48 hours provides a fairly accurate picture of the environment, it is not truly continuous observation. To most accurately explain the phenomena described in this study, true minute by minute observation of interactions as they happen would be required. The ability to capture activity beyond the first 30 reviews could provide a better view of the dynamics over a product's life cycle. Future research could consider technological data collection methods that could allow for more frequent observations of a larger selection of reviews.

Overall, this research demonstrates that both the context and content present in online review forums can have a substantial impact on CRI. The context of how and where consumers interact with reviews is shown to be influential through both the physical placement of reviews and the norms of the community they exist in. The content of the reviews as demonstrated by norm violations, is also shown to be influential in generating consumer-review interactions. Academics and managers alike need to consider both of these factors when evaluating interactivity within the dynamic review environment.

CHAPTER 3: NORMALIZING INTERNATIONAL ONLINE REVIEW CONTENT: THE CHARACTER COUNT LANGUAGE INDEX

Introduction

Global brand reputation is critical to multinational enterprise success (Chabowski, Samiee, and Hult 2013), and online consumer reviews can quickly damage – or enhance – brand reputation. Retail sites with online reviews facilitate the seeking and sharing of information about global brands. Online reviews are trusted by 70% of global consumers, making reviews the second most trusted form of marketing communications (Nielsen, 2012). These conversations matter, as word of mouth is more influential than advertising during key stages of product consideration (Bughin, Doogan, andVetvik 2010). Since online or electronic word of mouth (eWOM) is visible, enduring, and accessible (Duan, Gu, and Whinston 2008), it is strategically important for firms to collect, analyze and respond to online user-generated content (Yadav and Pavlou 2014). Firms seek to leverage the vast amount of digital information available in product reviews to gain insights about consumer attitudes and behavior. To create competitive advantage they need to collect and analyze consumer conversations in increasingly sophisticated ways (Netzer et al. 2012).

The analysis of reviews is now routine for single-country and single-language settings. The challenge for multinational retailers is to compare and interpret the flood of consumer online review content across country and language borders. This raises the question: *Does language affect online reviewing behavior?* To compare reviewing behavior across languages, multinational retailers need analysis methods that are scalable and meaningful. Differences in language however impede managers' ability to directly compare review content across countries. Before any meaningful analysis can take place, online review content needs to be normalized to account for differences in the languages in which they are written. This research seeks to provide a tool that can be used to normalize the effects of language in online review content allowing for more accurate and robust analysis of online reviews across countries.

The international context of online reviews is recognized as an area in need of additional research (King et al. 2014), especially as online customer conversations are increasingly taking place in languages other than English. It is estimated that only 55% of web content is in English (W3Techs 2014). The web enables consumers to become more international and cosmopolitan (Riefler, Diamantopoulos, and Siguaw 2012), while also maintaining distinctive language and country traits. Although brand-oriented consumers often have a global mindset, consumers use sites that cater to their particular language and country. Multinational retailers' country-specific and language-specific retail sites enable consumers to shop for and discuss global brands (such as Bosch, Canon, iPhone and Pampers) within a culturally familiar online environment. For example, as of 2015, Amazon has 14 country-specific sites covering nine languages, Walmart has eight country sites serving seven languages, and Alibaba-operated sites such as Taobao and Tmall are expanding to multiple languages.

Past research developed and used a variety of models and methods to examine online reviews (primarily in English). Review volume is a common first aspect to be analyzed, and insights can be gained from the quantity of online content. Volume is a valuable measure to firms, as it suggests the level of customer engagement. The volume of online user generated content can indicate consumer awareness, interest, and product knowledge (Chen et al. 2011). The quantity of web chatter has been linked to product sales (Dhar and Chang 2009) and stock performance (Tirunillai and Tellis 2012).

There are two main approaches to measuring review volume. One approach is to measure the number of discrete contributions, such as the number of reviews or ratings (Khare, Lebrecque, and Asare 2012). However, the quantity of reviews varies across product categories. For example, consumer electronics and books can be expected to attract more reviews than household appliances. Similarly, the quantity of reviews on a site is not simply a reflection of product interest; it is also a reflection of the level of customer traffic to the site. Well-established sites such as the U.S. Amazon site receive much higher traffic and many more reviews than the sites of smaller online retailers, or even newly-launched Amazon sites in other countries. This makes it problematic to use the number of reviews as basis for comparing review volume across product categories or across review sites.

A viable alternative for retailers is to use the length of online reviews as a measure of review volume. Many analyses of online reviewing behavior have utilized review length in their models (e.g., Pan and Zhang 2011). Table 6 provides examples of the significant research conducted on online word-of-mouth using review length measured by word count or character count. Although review length is an accepted metric for comparisons within a country site, inherent differences in language structure and characteristics have the potential to affect average review length. Unless language characteristics are taken into consideration, the use of review length can be misleading when comparing review volume across country sites. The challenge of assessing the degree to which language characteristics matter is that there is no accepted system or method to compare the length of online consumer word of mouth across languages.

Table 6. Selected publications using review length as a predictor variable

Article	Sample Language	Dependent Variable	Data Source
Chevalier & Mayzlin, 2006	English	Sales	Amazon.com & bn.com
Mudambi & Schuff, 2010	English	Review Helpfulness	Amazon.com
Archak, Ghose & Ipeirotis, 2011	English	Consumer Preference	Amazon.com
Ghose & Ipeirotis, 2011	English	Helpfulness & Sales	Amazon.com
Pan & Zhang, 2011	English	Review Helpfulness	Amazon.com
Sridhar & Srinivasan, 2012	English	Rating Valence	Independent Travel Website
Wu, 2013	English	Review Helpfulness	Amazon.com
Yin, Bond & Zhang, 2014	English	Review Helpfulness	Experimental & Yahoo Shopping

This research seeks to fill this gap through the construction of an index that facilitates the comparison of consumer online verbosity across six of the commonly used languages in an online retailing context: English, Chinese, Japanese, Spanish, French, and German. These languages represent almost two billion speakers worldwide (Ethnologue 2014), and account for more than three-quarters of global web content (W3Techs 2014). Because the average length of online reviews measures the verbosity of consumer word of mouth and quantifies the level of online consumer engagement, the newly developed index enables researchers, multinational retailers, and global brands to more accurately compare the quantity of word of mouth across languages and country borders. This index can also be used to ascertain if those differences in length due to differences in language structure indicate statistically and practically significant differences in consumer review verbosity. This is managerially relevant to multinational firms seeking to compare online engagement of consumers in multiple countries.

Theoretical Foundation of Language and Online Reviews

Verbosity

Past research has established review length as a simple but important measure of the quantity of customer word of mouth. Firms care about the length of reviews as an indication of consumer interest, engagement, or purchase involvement (Bughin et al. 2010; Chen et al. 2011; Liu 2006). Volume can serve not only as a sign of consumer awareness and interest, (Chen et al. 2011), but also as a proxy for quantity of information (Yin, Bond, and Zhang 2014). Review length can also denote the degree of customer verbosity, or wordiness.

Information quantity can enhance consumer decision-making. More verbose reviews are more diagnostic and more helpful to consumers (Mudambi and Schuff 2010; Pan and Zhang 2011; Schindler and Bickart 2012). As a result, retail sites often encourage consumers to post longer, more expository reviews. Chevalier and Mayzlin (2006) propose that longer reviews indicate greater reviewer interest (i.e., engagement), as they can contain more mixed positive and negative product information. Yin et al. (2014) made a similar observation, implying that longer reviews are more complete and reflect greater depth of content. Longer reviews can include information reflecting both positive and negative product attributes.

Character count has come to dominate word count as the dominant metric in many digital applications, including Twitter and Google Adwords. The rise of Twitter and other character-count based social media has increased attention to structural differences in languages. It has been established that a 140 character Tweet in Japanese or Chinese conveys more information than a 140 character Tweet in English (Rosen 2012). Yet there is no generally accepted explanation for these differences, and no established standard for comparing length of expression across languages. This is managerially important, as a difference in average character count between reviews in two languages, for example, German and Japanese, could indicate a difference in consumer engagement between the two countries, but this can be established only after taking into account the nature of the language in which the review was written.

Language Characteristics

Previous efforts to quantify differences between languages in the business environment provide important insights on broad categories, families, and branches of language. Language is a key element of the psychic distance measure that has been used to identify differences between countries and how they influence internationalization and information flow in global markets (Johanson and Vahlne 1977). Dow and Karunaratna (2006) broke languages into families, branches, and sub-branches, to establish the distance between languages and their influence on trade. Language also can inhibit communication across borders (Chen et al. 2006). Chiswick et al. (2005), for example, created a scale of linguistic distance based on the relative difficulty to learn a language. Lohman (2011) and Joshi and Lahiri (2012) developed a language barrier index that quantifies the structural features of languages by analyzing 192 linguistic features of 2,678 languages to enable the comparison of their relative similarity. While these measures provide useful tools to compare how similar or different languages are from each other, they provide little insight into how verbosity or the volume of content systematically varies across languages.

Past research indicates basic structural linguistic differences in the length of expression (Cosnier, Dols, and Fernandez 1986), or the volume of characters required to convey the same information. As Usunier and Roulin (2010: 201) note, "Many Asian languages use no gender, little or no personal pronouns, do not conjugate verbs and provide locutors with a relatively undersignified text, which requires much information from the context for the message to be understood by the receiver." The interpretation of content equivalence depends on whether an information-theoretic perspective is taken (Liao 2013). Neubig and Duh (2013) found that although languages with large character sets tend to have more information per character, this does not imply that the average Tweet in those languages conveys more meaningful information.

Researchers have attempted to create a standard of length comparison for Twitter. For example, Rosen (2012) pointed out that a 140 character Tweet in Japanese or Chinese conveys more information than a 140 character Tweet in English, Summers (2010) estimated 140 Japanese characters to be equivalent to 260 English characters (a factor of 1.8), while Liao (2013) identified an adjustment closer to 188 English characters (a factor of 1.34). Other estimates have also been offered, but without a consensus on the comparison method or benchmark. These efforts highlight the need for cross-language length comparisons. Standards may still be platform-specific, as the findings for Twitter may not be generalizable to other online platforms. In fact, Liao (2013) found that length and style of expression varies across Twitter, Wikipedia, legal documents, and TED talks. This implies a conversion rate for Twitter would likely not be appropriate for online reviews, as the style of expression is quite different.

Comparing the character count of online reviews across a multinational retailer's country sites can generate insights for researchers interested in the influence of language structure on eWOM volume, and can provide practical guidance to firms. The following section explains the development of an index that allows for the comparison of online review verbosity across languages. The index uses English as the baseline for comparison, with index scores for five other countries. This assumes English is the "home language" for the multinational retailer. However, this technique can be adapted to use other languages as a baseline. This is a practical asset that can benefit both retailers and academic researchers.

Research Method

The objective of this research is to develop a Character Count Language Index (CCLI) that acknowledges the potential influence of inherent language characteristics on review length, and provide a tool that enables research and analysis of online reviews across countries and languages. To develop the index, a sample of reviews was taken from six country sites of Amazon.com: the United States, France, Germany, Japan, China, and Spain. Since reviewing behavior varies between classes and types of products (Sen and Lerman 2007), reviews were examined in three product categories: books, cameras and MP3 players. To facilitate comparison, ten reviews of a book, camera, and an iPod were selected from each country site for analysis. A summary of the products used in the data collection is provided in Appendix 2.

Text of 30 reviews was collected from each Amazon site in the native language of the country, and computed the character count for each review. The reviews from each country were translated into the other five languages used in this study, and the character count was computed for each of the translated reviews. The resulting dataset contained character counts for 180 reviews in each language. For example, the English character counts consist of the 30 reviews collected in English plus 30 reviews from each of the other five languages translated into English.

Google Translate was used in the analysis, as it has been found to provide accurate and reliable translations (Aiken et al. 2011; NIST 2006). Although Google Translate was expected to perform well for the intended purpose of comparing character count, a supplementary analysis was performed to verify the efficacy of Google Translate for character count comparison. Bilingual and native speakers were hired to translate a sample of the native language reviews used in the study to English. The character counts of the translations from Google Translate to the manual translations by bilingual and native speakers were then compared. In each case, there was no significant difference in the character counts between the Google Translation and the human translation. Additional detail on the procedure and the results of that analysis are provided in Appendix 3.

The development of the index focused on the inherent characteristics of language, rather than on sociocultural norms for language expression. By using character counts of the home country reviews as well as foreign language reviews translated back to the native language of each home country, differences in the structural characteristics of written language are identified.

Results

Paired t-tests were used to determine if there are significant differences between the character counts of the same reviews across languages. The results of these tests are summarized in Table 7. The character counts of all language pairs included in the sample were significantly different from each other. The implication of this finding is that it is problematic to simply and directly compare the length of reviews written in different languages. Language affects the length of online reviews.

Table 7. Paired t-tests of language character counts included in the CCLI

Language Pair	Mean Difference	Std. Error	t-value	Sig.
English – French	-178.96	18.23	-9.82	.000
English – German	-146.90	13.39	-10.97	.000
English – Japanese	628.51	53.83	11.68	.000
English – Chinese	825.04	71.92	11.47	.000
English - Spanish	-162.32	16.21	-10.02	.000
French – German	32.06	6.78	4.73	.000
French – Japanese	807.47	71.17	11.35	.000
French – Chinese	1004.00	89.35	11.24	.000
French - Spanish	16.64	7.04	2.36	.019
German – Japanese	775.41	66.79	11.61	.000
German – Chinese	971.94	84.93	11.45	.000
German – Spanish	-15.42	7.26	-2.12	.035
Japanese – Chinese	196.53	18.42	10.67	.000
Japanese – Spanish	790.83	68.57	-11.53	.000
Chinese – Spanish	-987.36	86.59	-11.40	.000

N=180

These aggregate character counts for the 180 reviews in each language were used to create the Character Count Language Index (CCLI). The aggregate character count of the original and translated reviews for each language was divided by the aggregate character count for each of the other languages to determine the index value for each language pair. The resulting CCLI, shown in Table 8, allows for the comparison of review length across languages. For example, the first row in Table 8 illustrates the number of characters in French, German, Japanese, Chinese, and Spanish that it would take to equal 100 characters of English content. More simply, the CCLI figure for each of the other languages shows the comparable character count if a review in that language were in English. For example, the French index value of 114 indicates that it takes 14% more characters to express the same message in French than it does in English; a 100 character review in English would be 114 characters if the same review were written in

French. Similarly, if an English review is 100 characters, it would be only 50 characters long in Japanese. The index can be similarly used to compare any of the languages included in the sample. For example, a 100 character Chinese review would take 335 characters to write in Spanish. The CCLI provides a simple, useful, and accurate way to compare the relative length of consumer online reviews across languages, and can be used to provide insight into consumer engagement.

Table 8. Character Count Language Index

Original Language	English Translation	French Translation	German Translation	Japanese Translation	Chinese Translation	Spanish Translation
English	**100**	**114**	**112**	**50**	**34**	**113**
French	87	100	98	43	30	99
German	89	102	100	44	30	101
Japanese	202	231	226	100	68	228
Chinese	296	339	331	147	100	335
Spanish	88	101	99	44	30	100

N=180

Application and Discussion

The length of consumer online reviews is a common measure of consumer interest and engagement. This research establishes that structural characteristics of languages significantly affect the length of consumer online reviews. To accurately compare review length across language-specific websites, language structure needs to be taken into account. The Character Count Language Index (CCLI) developed in this study provides a methodological improvement for examining the volume of online word of mouth across languages. The index provides a method to acknowledge structural differences in languages. Use of the index can lead to a more accurate indication of the relative verbosity of consumer reviews across country sites.

This index has implications for a fast growing body of research on online reviews and consumer reviewing behavior. Many academic studies of online reviews have included review length in their models. Table 6 provides an example of the significant research done on online word-of-mouth using review length. Although most research to date on consumer user generated content has been examined in a single country context, future research is likely to take a multinational perspective. Extending past research to a multinational context will be limited unless the analysis corrects for structural language differences. The CCLI enables both an extension of prior research and an opportunity to expand the scope of research across countries.

An Application

To demonstrate how using the CCLI can impact the understanding of the relative engagement across country sites, the index was applied to a large set of online reviews collected from Amazon country sites in multiple languages. An existing set of 28,159 online reviews across a broad selection of product categories collected from Amazon sites in the United States, France, Germany, and Japan was used. The raw mean character counts of the reviews from each country were calculated. The character counts were then adjusted by the CCLI. The raw and adjusted character counts are shown in Table 9.

Table 9. Comparison of Raw and Adjusted Characters Counts Across Languages

		Raw Character Count		CCLI Adjusted Count	
Country	**N**	**Mean**	**SD**	**Mean**	**SD**
France	1,864	409.42	438.79	359.14	384.90
Germany	3,690	784.29	969.43	700.26	865.56
US	17,181	734.75	955.58	734.75	955.58
Japan	5,424	262.15	188.58	524.30	377.16

When looking at the raw character counts for each country, Japanese reviews are shortest and German reviews are longest. After adjusting for length using the CCLI, French reviews are shortest and the U.S. reviews are longest. Adjusting the review length by the index gives a more accurate view of consumer verbosity and product engagement. Without accounting for these differences, one might conclude that Japanese consumers are the least engaged at the level of the individual review, when in fact it is the French who display the lowest levels of engagement as measured by verbosity. This more accurate measure of verbosity can better inform multinational online retailers that seek to increase consumer engagement, as it shows where best to direct efforts to enhance engagement.

Future Research

As illustrated by the prior research highlighted in Table 6, review length is commonly included in empirical models. To extend these and other models to reviews in an international context, a method for adjusting for language is needed. The demonstrated application of the CCLI to a set of 28,000 Amazon reviews shows how the naïve, or unadjusted analysis yields substantially different, and potentially misleading results.

Using the CCLI will enable researchers to separate the effects of language structure from other country and culture-level effects that may impact consumer engagement, as measured by review verbosity.

By using character counts of reviews in a base language and the character counts of reviews in other languages translated back to the base language, differences in language structure that could affect review length are controlled for. As has been demonstrated, the CCLI can be used to evaluate the impact of language on review length across languages, and provides a methodology that can also be utilized to extend the CCLI to other languages. This is of practical use to multinational retailers, as they continue to expand the number of country-specific sites. Confirming and quantifying an inherent linguistic differential across languages in the context of online reviews can enable meaningful multinational comparisons of review length as a proxy for the level of consumer interest and involvement. Review character counts can now be used more effectively in cross-language comparison since the uncertainty of language structure differences can be taken into account.

For researchers, this also opens up new avenues for theory development regarding the effects of language and culture on business-related communications. While this research clearly demonstrates that language structure itself can influence the length of online WOM, it is likely that language structure alone is not the only variable responsible for variation in online WOM volume between countries. Cultural differences are also likely to play a role in reviewing behavior. Many researchers have distinguished clear differences in national culture (e.g., Hall 1976; Hofstede 1980) and these differences have been linked to culture specific consumer behavior and communication styles (Farley and Lehmann 1994; Steenkamp, Batra, and 2003). National culture and communication styles may not only influence the length or volume of online WOM, but potentially other characteristics like valence. Since country, culture, and language are often intertwined (Brannen 2004), this study takes a step toward untangling language from the equation, by providing a tool to address the inherent language characteristics that affect the volume of online WOM. Future research can build on the findings of this research to develop finer-grained explanations of differences in consumer culture and reviewing behavior.

As user-generated content becomes a more influential component of the global consumer's decision making progress, multinational firms recognize the challenge involved in analyzing the data available to them. Given the volume of data available, few firms have the tools and resources to painstakingly translate and compare product reviews across all country sites. There is a need to take a more systematic, strategic and scalable approach to analyzing reviews and social media content across country borders, and the CCLI provides firms with a practical tool for a more systematic approach. Taking the relative length of language into account is an essential component of accurately evaluating consumer reviews across countries. This study helps clarify the understanding of online word of mouth volume across countries and languages, and provides findings that are conceptually insightful and practically significant to international marketing strategy.

CHAPTER 4: THE INFLUENCE OF LANGUAGE AND CULTURE ON ONLINE REVIEWS

Introduction

Despite early expectations of the web enabling borderless commerce, language and country barriers continue to affect multinational retailers operating online. Consumers praise and disparage global brands using star ratings and comments in multiple languages through country-specific review sites. Firms are recognizing the power of online WOM, and gain a competitive edge by collecting and analyzing consumer conversations (Netzer et al. 2012). For retailer and consumer product firms, "listening in" effectively to consumers is essential business strategy (Landsman 2013). Online retailers such as Taobao, Testbericht, and Amazon actively solicit and promote consumer reviews on global brands. The benefits of reviews to online retailers include frequent and lengthier site visits, higher customer satisfaction, and higher purchase activity (Chevalier and Mayzlin 2006; Zhu and Zhang 2010). In addition, firms have the opportunity to gain key insights about consumer attitudes, social cognition, and decision processes from the star ratings, review content, and consumer interactions with reviews exhibited on the site.

Online consumer reviews are a prominent vehicle for consumers to share their experiences with global brands. Despite the volume of online reviews worldwide, cross-cultural aspects of eWOM has been identified as an area in need of further academic research (King, Racherla, and Bush 2014). While the web is borderless, consumers are not. Review sites facilitate the seeking and sharing information about brands. Consumers post reviews of global brands through country-specific and language-specific sites. Although studies on multinational aspects of e-commerce provide key lessons (Lim et al. 2004; Sia et al. 2009), cross-border online consumer behavior remains understudied. This research seeks to answer the question: How do language and sociocultural norms help explain consumer reviewing behavior and consumer interactions with online reviews across countries? This chapter develops and tests theory-grounded hypotheses in an analysis of 32,000 reviews of 187 products on Amazon.com retail sites in the United States, the United Kingdom, Germany, France, and Japan.

First, high-context/low-context cultural perspectives are examined to test assumptions regarding WOM volume across countries. The results find that after adjusting for language structure, reviews from high-context countries are not necessarily less verbose than reviews in low-context countries. Second, past insights on cross-cultural bias is incorporated in an analysis of WOM valence. Assessing the valence of text comments is inherently difficult, even for a single language, and poses special challenges for cross-language evaluations. To avoid making misleading comparisons of consumer comments, the investigation of valence is limited to review star ratings. Although analyzing the valence of star ratings is more straightforward than analyzing translated text, potential biases remain. Extreme response bias and negative response bias are identified as potential cross-cultural biases in online reviewing behaviors that affect star rating valence. The results find that online reviews in languages in countries associated with weak extreme response bias and strong negative response bias to have a lower mean star rating than in other countries. Acquiescence bias and uncertainty avoidance are also examined in the context of voting on the helpfulness of reviews. The results indicate a higher percentage of helpful reviews in countries associated with acquiescence bias and that review extremity is considered more helpful in countries with high levels of uncertainty avoidance.

Prior Research on Online Reviews

Firms recognize that consumers are increasingly influenced not only by marketing communications and product attributes, but also by what other consumers say about the brand. Social cognition, defined as how individuals think about others, affects consumer decisions and provides a rationale for the importance of consumer WOM to firms. According to social comparison theory (Festinger 1954), individuals have a basic drive to evaluate their own opinions by looking to others for social cues that signal taste and values. This explains why consumers find information personally delivered from a non-marketer to be highly credible (Herr Kardes, and Kim 1991).

Social comparisons are made online and offline. Online, open-ended peer comments can emulate the subjective and social norms of offline interpersonal interaction (Kumar and Benbasat 2006), and consumption-related online communities can act as a strong social network for consumers (Brown, Broderick, and Lee 2007; Schau, Muñiz, and Arnould 2009). The conceptual framework of service dominant logic can also explain the influence and importance of consumer WOM, as review sites enable consumers and firms to co-create brand meaning and value (Vargo and Lusch 2004). In the digital environment, firms do not have full control of brands, as consumers have the power to shape brand image and reputation. Since online WOM is visible, enduring, and accessible (Duan, Gu, and Whinston 2008), it is strategically important to firms to collect, analyze and respond to online user generated content.

Past research on online reviews has emphasized attributes such as volume and valence (Floyd et al. 2014). Volume is a measure of WOM quantity and is an indication of consumer awareness, interest, and knowledge (Chen, Wang, and Xie 2011; Liu 2006). Volume is also used by practitioners to assess WOM impact (Bughin et al. 2010). Although past research has not provided the granularity to compare the volume or verbosity of discussion of online content across languages, measures of the quantity of online content are valuable to firms. Firms have linked the quantity of web chatter to product sales (Dhar and Chang 2009) and stock performance (Tirunillai and Tellis 2012). The volume of user generated content has been measured as the number of contributions, such as the number of reviews or ratings (Khare, Lebrecque, and Asare 2012), and as the size of the contribution, such as word or character count (Mudambi and Schuff, 2010). Past research has established review length as a simple but important measure of the quantity of user generated content. Firms care about the length of reviews as an indication of consumer interest, engagement, or purchase involvement (Bughin et al. 2010; Chen et al. 2011; Liu 2006). Longer reviews are more diagnostic and helpful to consumer decision making (Mudambi and Schuff 2010; Pan and Zhang 2011; Schindler and Bickart 2012). As a result, retail sites often encourage consumers to post longer, more expository reviews. Chevalier and Mayzlin (2006) propose that longer reviews indicate greater reviewer interest (i.e., engagement), as they contain more mixed positive and negative product information. Yin et al. (2014) made a similar observation, implying that longer reviews are more complete and reflect greater depth of content. Longer reviews can include information on both positive and negative attributes.

Valence indicates whether sentiment is positive or negative. Valence is reflected in the content of text comments and by the number of “stars” reviewers give the product. The valence of consumer reviews significantly affects consumer attitudes about the product, influences consumer purchase decisions (Chevalier and Mayzlin 2006), and indirectly affects sales (Duan et al. 2008; Floyd et al. 2014). In addition to the directional importance of review valence, the variance of these star ratings is also meaningful. A high standard deviation of product ratings has been shown to improve sales for products with low aggregate ratings (Sun 2012).

The perceived helpfulness of a review has also been a subject of considerable research. Many review sites prompt consumers to indicate whether or not they think a review is helpful or not. Amazon for example poses the question “Was this review helpful to you?” for each review. The helpfulness of a review could reflect the diagnosticity of a review, and a website with more helpful reviews can offer more value to consumers (Mudambi and Schuff, 2010). Please refer to Table 10 for highlights of key literature concerning review volume, valence, and helpfulness.

Table 10. Key Findings on Review Volume, Valence, and Helpfulness

Study	Language of Reviews used in Study	Key Findings
	Review Volume	
Liu (2006)	English	Word of mouth can be used to explain movie box office revenue, but only in terms of volume, not valence.
Duan, Gu, and Whinston (2008)	English	Accounting for endogeneity, the valence of online reviews does not impact movie box office revenue, but volume does.
Chintagunta, Gopinath, and Venkataraman (2010)	English	Controlling for other factors, the main driver of movie box office revenue is the volume of online reviews.
	Review Valence	
Sen and Lerman (2007)	English	Product type moderates the effect of review valence; consumers have a negativity bias for utilitarian products.
Kim and Gupta (2012)	English	Negative emotional expressions in a review decrease its' informative value and influences consumers product evaluation
Qui, Pang, and Lim (2012)	English	Review valence moderates the effect conflicting aggregate ratings have on perceived review credibility and diagnosticity
	Review Helpfulness	
Mudambi and Schuff (2010)	English	Review extremity, review depth, and product type affect the perceived helpfulness of reviews.
Ngo-Ye and Sinha (2014)	English	The recency, frequency, and monetary value of a reviewers transaction history drives the helpfulness of their reviews.
Yin, Bond, and Zhang (2014)	English	Online review containing content associated with anxiety were found to be more helpful than reviews with content indicative of anger

The assessment of consumer opinions on global brands in the online environment requires the interpretation of WOM across countries and languages. Despite a large body of literature on online WOM, there has been little cross-language or cross-country research on online consumer reviews. This is an important research gap. Examining the volume, valence, and helpfulness of consumer reviews across country and language boundaries can challenge both theoretical and practical assumptions underlying retailing.

Overall, this study represents a conceptual contribution that involves “envisioning a new reality” (MacInnis 2011) in consumer online word-of-mouth by examining the under-explored dimension of cross-language differences in WOM communication. The study identifies and measures differences in how consumers express themselves across language-specific country websites. Results indicate that large-scale analyses of WOM across countries need to account for language differences in the volume of communication, potential cultural response biases in the use of standard 5-star rating systems and consumer perceptions of review helpfulness. This study builds on a rich and diverse theoretical foundation, as summarized in the following section.

Conceptual Framework and Hypotheses

Language in International Marketing

Language influences thinking in direct and indirect ways (Hunt and Agnoli 1991; Whorf et al. 2012), resulting in differences in how people express themselves and conceptualize the world around them. Past research on language within the context of international business has generated insights regarding the MNE as a multilingual community (Luo and Shenkar 2006). Much of the research has assumed that language is a barrier, rather than an opportunity (Chidlow, Plakoyiannaki, and Welch 2014).

While it is clear that language is of critical concern to multinational marketers, the influence of language on consumer WOM has remained largely unexplored. Language can affect the behavior and decision making of customers in the retail context. However, since country, culture, language, and market factors are intertwined (Brannen 2004), disentangling language from culture and market factors is difficult. Past research in marketing has established the interconnected nature of country, culture, consumer attitudes toward global brands, and choice and purchasing behavior (Farley and Lehmann 1994; Steenkamp, Batra, and Alden 2003). Differences in consumer needs, preferences, and behavior have also been documented (Steenkamp, Batra, and Alden 2003), with the variation explained by market variables and product attributes, as well as by cultural values and norms (Dawar, Parker, and Price 1996; Schwartz 1999). To explain online WOM more specifically, it is important to look to aspects beyond consumer access to the Internet, and systematically examine the role of language and culture. To build upon the current online review literature, this research specifically focuses on the potential influences of language and culture on the volume, valence, and helpfulness of consumer WOM.

Cross-Language Verbosity of Discussion

The volume of consumer online discussion has been established as an indication of consumer awareness, interest, and knowledge (Chen et al. 2011), and as a proxy for information (Yin, Bond, and Zhang 2014). However, theories on communication styles and language structure indicate that measuring WOM volume as word count or as character count can lead to misleading comparisons across languages. A difference in average character count of German reviews and Japanese reviews could indicate a difference in consumer engagement, but this becomes clear only when communication styles and language structure are taken into account.

High-context and Low-context Communication.

Hall's (1976) categorization of communication styles as high-context or low-context has been widely adopted as a lens to analyze communication through the influence of a country's culture, especially for oral, face-to-face conversations. High-context cultures place more emphasis on non-verbal communication and symbolic behavior expressed through body language, silence, and proximity (Wurtz 2005). Communication in high-context cultures is also less direct, as messages are conveyed more through context than words (Kim, Pan, and Park 1998; Usunier and Roulin 2010; Wurtz 2005). High-context communicators prefer more understated communications (Gudykunst, et al. 1996). China, Japan and most Asian cultures are generally considered to be high-context (Hall 1976). Unlike most Western countries, France has been described as a semi-high context culture (Usunier and Roulin 2010) or a medium context culture (Singh, Kumar, and Baack 2005).

In contrast, low-context cultures such as Germany, the US and the UK emphasize direct communication and place more value on verbalization than on social context. The meaning of low-context messages are transmitted primarily by the information contained in the message itself (Hall 1976). These differences in communication style influence the openness and dramatics of communications. de Mooij and Hofstede (2011) found that individuals in low-context cultures are more verbally-oriented than their high-context counterparts. Communication in a low-context culture can produce more colorful words, exaggerated stories, and picturesque speech than communications in a high-context culture (Gudykunst et al. 1996).

Hall's framework suggests a culture's communication style can influence the volume and nature of online consumer communications. Communication style differences have been shown to matter in Internet-based communications (Usunier and Roulin 2010). Extensions of Hall's (1976) framework have been used to explain the amount of text and illustration used on websites, the directness of sales approaches employed, and the perception of website effectiveness (Hermeking 2005; Singh et al. 2005; Wurtz 2005). Communication style also influences interactivity. Consumers from low-context cultures do more searching, editing, and manipulating of internet content (Cho and Cheon 2005). Marketers must pay attention and respond to these cultural differences (Tian and Emery 2002) and design websites and communication forums accordingly. For example, Burgman et al. (2006) found cultural dimensions can influence the symmetry, focus, and number of hyperlinks on a website, as well as attitude toward a website, navigational ease, and purchase intent (Baack and Singh 2007).While extensive work has examined the influence of high and low-context communication on website design, the impact of communication styles on the volume of consumer online expression has not been established.

Communication context is expected to affect the verbosity of written discussion in online consumer reviews. Communication style should be strongly related to differences in verbosity of reviews across countries, so reviews from countries associated with high-context cultures will have lower levels of verbosity, as measured traditionally by character count, than reviews written from countries associated with low-context cultures. Therefore, it is hypothesized:

H1: The more a language is associated with a high-context culture, the lower the verbosity of online reviews in that language.

Cultural Response Bias in Reviews

Research on cultural response bias indicates how culture affects consumer expression of attitudes and opinions. Cultural response bias is evident when a respondent's systematic way of answering represents typical behavioral characteristics of a culture, rather than a relationship to the question content or type of item (Roster, Albaum, and Rogers 2006). In the context of online content, of interest are underlying cultural tendencies and the way they are revealed using measurement scales.

One widely acknowledged cultural response bias is the Western tendency to explain or attribute causality to inherent disposition or nature, and the Eastern tendency to attribute causality to the situation or context (Choi, Nisbett, and Norenzayan 1999). Psychology theory suggests that East Asians are likely to take a holistic approach, and judge a person or product in a specific situation, while Westerners take a more analytic approach, categorizing specific attributes (Nisbett et al. 2001). Westerners synthesize evidence, perhaps by ignoring situational factors, while East Asians judge a product in the specific context or situation (Choi, Nisbett, and Norenzayan 1999). In the product review context, theory suggests that an East Asian reviewer may like and use a brand, but more harshly evaluate a new model in a specific situation or context.

Several types of cultural response bias have been identified and examined in questionnaires translated into multiple languages. Country-level variation in opinion-seeking and opinion-giving behavior have been also been found (Dawar et al. 1996). Differences in consumer responses may be due to true consumer differences between countries regarding the underlying construct, or may be due to systematic biases in the way people of different countries respond to certain items (Steenkamp and Baumgartner 1998: 78). Cultural response bias has explained differences in consumer responses at the country level (Baumgartner and Steenkamp 2001; Steenkamp and Ter Hofstede 2002; Tellis and Chandrasekaran 2010). Measurement invariance and a lack of cross-national construct equivalence are important considerations (Diamantopolous and Papadopolos 2010; Hult et al. 2008; Steenkamp and Baumgartner 1998).

Past research has generally examined response style across multiple instruments that utilize scaled items of varied range and complexity. In contrast, the language, instrument, scales, and process of a typical online review are much less involved than an in-depth survey instrument. Online review sites often rely on one scaled item, a product rating of one to five stars. Previous research has not examined cultural response bias in this simple environment.

For reviews, of particular relevance is extreme response bias (Van Herk, Poortinga, and Verhallen 2004). This is the tendency to endorse extreme response categories on a rating scale, regardless of content. On review sites, one-star and five-star ratings can indicate attitude extremity, defined as the distance from the midpoint of an attitude scale (Krosnick et al. 1993), with a midpoint rating interpreted as a moderate view. This can be influenced by personality, culture and language. According to Roster et al. (2006), culturally-based boundaries can restrict the valence and intensity of attitudinal expressions within normative bounds. Past studies found some countries are more likely to have an extreme response bias, but results are mixed. While Tellis and Chandrasekaran (2010) found that Asian respondents utilize the extreme end of the spectrum, several other studies on the cross-cultural equivalency of Likert scales indicated that consumers in Asian countries value moderation and are less likely than US consumers to use the extreme points of one and five (Johnson et al. 2005). Considerable research has found that online reviews are generally positive, with one star reviews rare (Hu, Zhang, and Pavlou 2009). Thus, a tendency to avoid extreme responses can imply that positive reviewers in Asian languages are less likely to use five stars than reviewers in other languages.

A related response style is negative response bias or "nay-saying." Research in this area indicates a clear language predisposition. The US has a weak tendency for a negative response, in contrast to the stronger negative response bias of other countries. In particular, Tellis and Chandrasekaran (2010) found a strong Japanese tendency of "nay-saying" in written responses. This implies that countries or languages with a negative response bias will rate products lower than will others. Taken together, the extreme response bias and the negative response bias work in the same direction. Countries such as Japan have a weak extreme response bias and a strong negative response bias, and other countries (e.g., the US) have a strong extreme response bias and a weak negative response bias.

Star ratings appear simple and straightforward, and so may not exhibit the response bias found in other contexts. As a result, it is understandable if retailers attempt to gauge relative consumer sentiment and approval by comparing the star rating of a global brand across multiple countries. However, since theory and past research have indicated cross-cultural differences in scale utilization in other contexts, a 4-star rating may not have the same meaning to a customer in one country as a 4-star rating to a customer in another country. Theoretical arguments on holistic versus analytic thinking styles, extreme response bias and negative response bias, all lead to an expectation of valence differences across languages. For online reviews, it is hypothesized that:

H2: Cultural response biases affect online review valence, reflected in the star ratings in that culture's reviews.

More specifically, online reviews in languages in countries associated with weak extreme response bias and strong negative response bias have a lower mean star rating than in other countries. Since Japan is associated with a holistic and situational perspective, a weak extreme response bias and a strong negative response bias, Japanese reviews are expected to have a lower mean star rating than reviews in the US, UK, France and Germany.

Acquiescence Bias and Collectivism

There is more to consumer reviewing behavior than the volume and valence of posted reviews. Consumers also decide whether or not to visibly engage with the content posted by other consumers. This relates to theory on the cross-cultural response bias of acquiescence or agreement bias. Acquiescence bias is the tendency to agree or strongly agree with scaled items, regardless of item content (Van Herk, Poortinga, and Verhallen 2004). This tendency is also known as "yea saying." Interestingly, this can co-exist alongside a tendency towards "nay saying" in a different context (Tellis and Chandrasekaran 2010). Although Japanese may have a nay-saying bias in some written contexts, the use of language in Japan encourages the use of yes in interactive or interpersonal situations, even where the speaker means "no" (Tannen 1984). This response bias may also be partially explained by the dialectical East Asian self-concept and cognitive approach, which is more holistic and tolerant of contradiction than in Western cultures (Spencer-Rodgers et al. 2004; Spencer-Rodgers, Williams and Peng 2010).

Research has shown that the general tendency to agree or disagree is culturally related (Van Herk et al. 2004), and is more common in Asian languages (Smith and Reynolds 2001). One of the most common ways to compare culture is Hofstede's (1980) measure of individualism and collectivism. These dimensions have an effect on communication because they affect the norms and rules that guide behavior within cultures (Gudykunst et al 1996). Low-context communication is typically associated with individualistic cultures, while high-context communication is typically associated with collectivist cultures (Gudykunst and Ting-Toomey 1988). Members of individualistic cultures place value on independence and individual achievement, while members of collectivist cultures value on solidarity and harmony (Hofstesde 1980). Individualistic cultures also stress individual goals while collectivist cultures are more concerned with the goals of the group as a whole (Ngai et al. 2007; Triandis 1988). Prior work has shown that acquiescence bias is positively associated with collectivist cultures (Johnson et al. 2005).

The Amazon site question of "Was this review helpful to you?" has the potential for acquiescence response bias. Consumers have two response options of yes or no. If a culture has a greater tendency for expressing agreement, this would be indicated by a greater prevalence of reviews evaluated as helpful through a greater number of "yes" responses to that question. A tendency to agree is likely to lead to a tendency to find a review helpful. Cultures higher on the dimensions of collectivism are more likely to demonstrate acquiescence bias in responses. Therefore, due to acquiescence response bias associated with culture it is expected that:

H3: Online reviews in countries associated with acquiescence bias will have a higher percentage of reviews voted helpful than reviews in other countries.

More specifically, since Japanese communications are associated with an acquiescence bias in interactive settings, and Japanese culture is higher in the dimensions of collectivism than their Western counterparts, it is expected that a higher percentage of Japanese reviews will receive a "yes" vote, indicating the review is helpful.

Uncertainty Avoidance

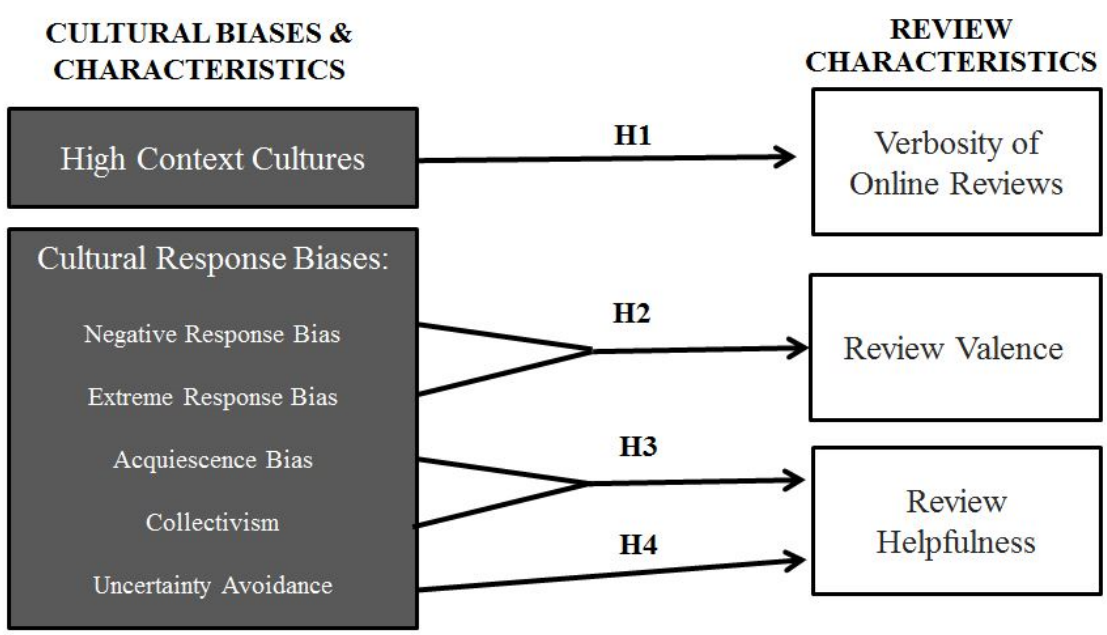
While it is expected that acquiescence bias and collectivism will lead to countries like Japan having a higher percentage of helpful votes, it is likely this relationship has an additional level of complexity. The valence of a review again needs to be considered when evaluating review helpfulness across cultures. Hofstede's (1980) cultural dimension of uncertainty avoidance helps to explain this relationship. Cultures with higher uncertainty avoidance have many rules and little tolerance for ambiguity (Hofstede 2001). These cultures also actively avoid risk to create a sense of control (Hofstede 1980). The uncertainty avoidance dimension has also shown to be useful in evaluating consumer behavior and internet usage. Low uncertainty avoidance shoppers are more open-minded when searching for information (Vishwanath 2003). Consumers in high uncertainty avoidance cultures use more WOM sources than their low uncertainty avoidance counterparts (Money, Gilly, and Graham 1998) suggesting that the use of online reviews, and their characteristics, may be even more critical for these consumers. Park and Lee (2009) demonstrate this effect by showing that Korean consumers (high uncertainty avoidance) use online reviews more frequently than American consumers (low uncertainty avoidance).

Shopping online is associated with many perceived risks (Forsythe and Shi 2003). Online reviews are used by consumers are a tool to reduce uncertainty and risk in purchase decisions (Zhu and Zhang 2010). When considering the valence of online reviews using the common one to five star scale, different points on the spectrum may be more useful in mitigating risk than others. While it has been shown that consumers find negative reviews to be more helpful (Mudambi and Schuff 2010), the consideration of consumers' national culture may influence this effect. Extreme reviews (one or five star) offer more certainty as to the evaluation of a product by a reviewer than a review with a star rating in the middle of the scale. Accordingly, review extremity may be more valuable and viewed as more helpful to a consumer from a high uncertainty avoidance country as they tend to avoid the ambiguity that may be associated with a review in the middle of the scale. Therefore it is hypothesized that:

H4: Extreme reviews will be considered more helpful than moderate reviews in countries associated with high uncertainty avoidance.

More specifically, it is expected that countries associated with high uncertainty avoidance will consider extreme reviews more helpful than moderate reviews. Since Japan is associated with a high level of uncertainty avoidance, it is expected that extreme Japanese reviews (one and five star reviews) will be found more helpful than moderate Japanese reviews (two, three, and four star reviews). Please refer to Figure 3 for a summary of the hypotheses and conceptual model.

Figure 3: A Conceptual Model of the Influence of Language and Culture on Online Review Content



Research Methods

Data Collection

The hypotheses are tested by analyzing online reviews across countries and languages. This study seeks to understand how and why consumer review activity differs across language and country boundaries by analyzing data from the world's most popular online multinational retailer, Amazon.com. The Amazon site is also the sixth most popular site globally across all site categories (Alexa Internet 2013). Although Amazon sells products from thousands of MNEs, each Amazon site is language and country-specific, a strategy common to many multinational retailers. Amazon's country sites are the most popular shopping sites in the US, UK, Japan, Germany and France, and the sites rank in the top five overall in terms of traffic and page views in each country (Alexa Internet 2013).

The Amazon Product Advertising API was used to collect every review of 187 products from Amazon sites in Japan, France, Germany, the U.K., and the US. A total of 32,556 reviews were collected. A software agent retrieved the review data directly from Amazon's web service and automatically coded additional data for each review. Thirty-one percent of the reviews came from the European sites, 17% from Japan, and 52% from the US. A diverse and representative set of product categories was chosen. Search goods in the sample included cameras, toasters, coffee makers, sandwich grills, baby gyms, cleaning products, and dental hygiene products. Experience goods included music CDs, fiction and nonfiction books. Since the same products were not always available in every country, and since product competition and consumer tastes vary, product selection in each country was determined by the availability of a large number of reviews. In each country, products that were heavily reviewed were selected. The set of products is similar, not identical, across countries due to market variation, but there is no indication of systemic differences between the sets of products on key measures. It is worth reemphasizing that the focus of this research is not to compare reviews at the product level. Instead, the motivation is to examine the implications of making country-to-country comparisons of verbosity, star ratings, and helpfulness.

To test hypotheses one and two, the full dataset of 32,556 reviews described above was used. The descriptive statistics for this dataset are displayed in Table 11. To test hypotheses three and four, reviews that did not have at least one vote on the question, “was this review helpful to you?” were excluded from the analysis. This resulted in a sample of 23,332 reviews. The descriptive statistics for this dataset are shown in Table 11. For clarity, a brief explanation of each variable used in the hypothesis testing is needed. Through the analyses, the variables raw character count, adjusted character count, star rating, and helpfulness are the variables used. Raw character count is the character count of each review in its’ native language. Adjusted character count takes the raw character count and applies the Character Count Language Index developed in Chapter 1, to normalize the character count of each review by accounting for structural language differences. Star rating is the star rating for each review based on Amazon’s 1-5 star rating system. Review helpfulness determined by an affirmative response to the question “Was this review helpful to you?” that Amazon poses for each review. The specific application of each of these variables for each hypothesis test is explained further in the following section.

Table 11. Chapter 4 Descriptive Statistics

	Full Sample											
	US		UK		Germany		France		Japan		Aggregate	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Raw Character Count	734.75	955.58	665.88	718.19	784.29	969.43	409.42	438.78	262.15	188.58	633.70	841.99
Adjusted Character Count	734.75	955.58	665.88	718.19	70.28	865.56	359.15	384.91	524.30	377.16	664.97	824.60
Star Rating	4.23	1.22	4.29	1.18	4.24	1.20	4.26	1.08	3.57	1.66	4.13	1.31
Helpful Percentage	.44	.44	.47	.44	.46	.41	.42	.44	.71	.31	.49	.43
Total Votes	5.55	36.65	7.80	30.29	10.44	35.00	4.23	10.24	53.95	97.317	14.4	53.578
N	17,181		4,397		3,690		1,864		5,424		32,556	

	Adjusted Sample – Reviews with no responses to the question "Was this review helpful to you?" removed											
	US		UK		Germany		France		Japan		Aggregate	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Raw Character Count	860.41	1085.97	766.04	800.58	880.75	1,073.63	467.38	499.35	264.65	190.05	700.41	933.61
Adjusted Character Count	860.41	1085.97	766.04	800.58	789.41	958.60	409.99	438.04	529.31	380.09	744.83	908.98
Star Rating	4.07	1.35	4.17	1.29	4.13	1.30	4.10	1.21	3.52	1.68	3.97	1.43
Helpful Percentage	.66	.38	.69	.36	.61	.36	.69	.36	.75	.27	.68	.35
Total Votes	8.42	44.87	11.52	63.22	13.99	39.90	6.98	12.40	56.89	99.10	20.09	62.38
N	11,326		2,978		2,755		1,130		5,143		23,332	

Analysis and Results

Comparing Cross–Language Verbosity

To test H1, the mean raw character count by country using the full data set of 32,556 reviews was analyzed. A regression model with raw (unadjusted) character count as the dependent variable and four independent binary dummy variables representing the United Kingdom, France, Germany, and Japan was used to conduct the analysis. Since Amazon is a US based multinational, the United States was the baseline. The result shows that there are significant differences ($p < 0.000$) in review length between the US and the other country sites in the sample. As expected, Japan had the lowest average character count (262.15), as it is most associated with high-context communications and culture. The reviews of the other sites, all in western countries with lower-context cultures, were more verbose (ranging from 409.42 for France to 784.29 for Germany). France, which has been described as medium-context (Singh et al 2005) or semi-high context (Usunier and Roulin 2010), was the least verbose of the Western countries analyzed. The results are illustrated in Table 12. These findings provide support for hypothesis 1.

Table 12. Comparison of Character Counts Across Country/Language (US as baseline, full sample)

Raw Character Counts

Language	Mean Raw Char. Count	Diff. from US English	SE	t-value	Sig.
English (US)	734.75				
English (UK)	665.88	-68.87	13.89	-4.96	0.000
French	409.42	-325.33	20.04	-16.24	0.000
German	784.29	49.55	14.90	3.32	0.001
Japanese	262.15	-472.60	12.80	-36.93	0.000

CCLI-Adjusted Character Counts

Language	Mean Adjusted Char. Count	Diff. from US English	SE	t-value	Sig.
English(US)	734.75				
English(UK)	665.88	-68.87	13.82	-4.98	0.000
French	359.15	-375.60	19.94	-18.84	0.000
German	700.26	-34.49	14.83	-2.32	0.020
Japanese	524.30	-210.45	12.74	-16.52	0.000

N=32,556

Examining the raw character counts alone however does not provide an adequate evaluation of cross national verbosity due to inherent difference in language structure that can influence length of expression. As discussed in the previous chapter, to compare the verbosity of online content, character counts first need to be adjusted to account for inherent language differences that can influence the relative verbosity of online reviews. Accordingly, to further test hypothesis 1, the Character Count Language Index (CCLI) developed in the previous chapter was applied to the data set to normalize the data and remove the potential influence of language structure on review length.

Hypothesis one was again tested using a similar regression model as used with the raw data, but this time using the CCLI-adjusted character count as the dependent variable (see Table 11). After the adjustment, the differences between the United States baseline and the other countries were again significant ($p < 0.02$ for Germany, $p < 0.000$ for the others), however the spread between the high-context and low-context countries was substantially lower. Having adjusted for language difference, the support for hypothesis one becomes less clear. While the results in Table 12 confirm that verbosity levels differ by country, even when using the CCLI-adjusted character counts, the relationship between context and verbosity is now less clear. With the adjustment, French reviews are less verbose than Japanese reviews (359.15 versus 524.30), and an additional t-test shows this difference to be significant ($p < 0.000$). This provides interesting evidence that although verbosity has been attributed to the high/low context, verbosity differences could actually be, in part, an artifact of differences in language structure. In other words, when taking language structure into account, the cultural communication context of the country does not fully explain differences in verbosity. More specifically, the structure of a language likely plays a role in determining the relative verbosity of online review content across countries, not just culture alone. This highlights a rich area for future research.

Testing Response Bias

Valence of reviews is the degree of positive consumer sentiment expressed about a product. In the case of Amazon.com, consumer reviewers provide a quantitative evaluation of sentiment in a rating from one to five stars, along with open ended text comments. To test H2 on cultural response bias in online reviews, the set of 32,556 reviews collected from Amazon sites in Japan, France, Germany, the U.K., and the US was used. For each review, the data included the star rating (valence), review length in characters, and CCLI adjusted characters. A regression model with star rating as the dependent variable and dummy variables for the US, UK, France, and Germany as the independent variables, with Japan as the baseline was used. It was hypothesized that there would be an effect of response bias on consumer reviewing behavior. It was expected that extreme response bias would affect online review valence. More specifically, reviews in languages with a weak extreme response bias and with a strong negative or "nay-saying" response bias (i.e., Japanese) will have a lower mean star rating. The regression results are shown in Table 13. Japanese reviews had a mean star rating of 3.57, and the other mean star ratings ranged from 4.23 for the US to 4.29 for the UK (see Table 12). Therefore, H2 was strongly supported ($p < 0.000$), providing strong evidence that the cultural response bias present in Japan does indeed lead to lower online review star ratings.

Table 13. Analysis of Valence Across Countries Regression Results: Star Rating Differences

Country	Descriptive Statistics			Country Differences (using Japan as a baseline)			
	N	Mean Star Rating	SD	Coefficient	SE	t-value	Sig.
France	1864	4.26	1.076	0.691	0.035	19.96	0.000
Germany	3690	4.24	1.202	0.674	0.028	24.51	0.000
UK	4397	4.29	1.177	0.725	0.026	27.72	0.000
US	17181	4.23	1.222	0.664	0.020	33.07	0.000
Japan	5424	3.57	1.66				
Total	**32555**	**4.13**	**1.31**				

$F = 313.82$
$p = .000$
$r^2 = 0.037$

Testing Acquiescence Bias

Consumers can engage with content posted by others by voting on the quality of the review, either agreeing or disagreeing that the review was helpful. Due to the role of acquiescence bias, Hypothesis three expected there to be a higher percentage of helpful reviews in countries associated with acquiescence bias, such as Japan.

To test hypothesis 3, review helpfulness was used as the dependent variable (with country dummy independent variables for the US, UK, France, and Germany, with Japan as the baseline). Reviews that did not have at least one vote on the question, "was this review helpful to you?" were excluded from the analysis. This resulted in a sample of 23,332 reviews. For Japanese reviews, 75.1% of the votes indicated the review was helpful. The helpfulness of reviews in other countries was significantly lower, ranging from 61% in Germany to 69.7% in the United Kingdom (see Table 14). Hypothesis three was strongly supported (p<0.000), indicating that acquiescence bias across countries makes a difference in the reported helpfulness of online reviews. This finding spurred a more general curiosity about potential differences in review helpfulness across countries and languages.

Table 14. Regression Results: Analysis of Helpfulness Across Countries (Comparison of Helpfulness Percentage)

Country	Descriptive Statistics			Country Differences (using Japan as a baseline)			
	N	**Mean Helpfulness Percentage**	**SD**	**Coefficient**	**SE**	**t-value**	**Sig.**
France	1130	0.687	0.357	-0.064	0.012	-5.57	0.000
Germany	2755	0.610	0.364	-0.141	0.008	-17.04	0.000
UK	2978	0.697	0.357	-0.054	0.008	-6.72	0.000
US	11326	0.660	0.377	-0.091	0.006	-15.46	0.000
Japan	5143	0.751	0.270				
Total	**23332**	**0.680**	**0.353**				

F = 91.38
p = .000
$r^2 = 0.015$

Testing Language and Cultural Effects on Review Helpfulness

To further analyze the influences of culture on review helpfulness a more nuanced approach is needed. Prior research has developed models to determine variables that are influential in determining review helpfulness in single country settings (see Mudambi and Schuff 2010). A goal of this research is to take prior research a step further by examining and comparing review helpfulness in a multi-country setting. Several factors, including valence and volume, have been shown to work together to influence review helpfulness in single country settings. By applying methods established in prior research to a multi-country analysis of review helpfulness, differences in the importance of variables that drive review helpfulness can be identified, and suggest underlying cultural influences that may account for these differences. Accordingly, a modified review helpfulness model was adopted from Mudambi and Schuff (2010) and applied to the data collected in this study to both test hypothesis four and determine the drivers of review helpfulness across countries.

To evaluate review helpfulness, the dependent variable in this analysis is Helpful Percent, which represents the percentage of affirmative responses to the question “Was this review helpful to you?” Rating and adjusted character count are included as independent variables. The Rating variable represents the star rating given to the review on the 1-5 Amazon rating scale; we added a squared Rating term to account for the expected non-linear relationship between valence and helpfulness in countries associated with high uncertainty avoidance. Adjusted Characters represents the character count of each review after being adjusted for language differences using the CCLI. This has been included as it is likely that longer reviews, which contain more information, may be viewed as more helpful. Total Votes is included as a control variable to account for the dependent variable being represented as a percentage. While prior research has analyzed product type as a potential influence in determining review helpfulness, it was not included in this analysis in an effort to isolate the results to reflecting differences between countries without the complication of product type. This results in the following model which is used to test the determinants of review helpfulness in this analysis:

$$HelpfulPct = \beta_1 Rating + \beta_2 Rating^2 + \beta_3 Adjusted\ Character\ Count + \beta_4 Total\ Votes + \varepsilon$$

The dataset of 23,332 reviews used to test hypothesis three was applied to the model. For ease of interpretation, the regression model was run separately for each of the five countries included in the sample (US, UK, France, Germany, and Japan). The results of the analysis, displayed in Table 15, yield interesting results.

Table 15. Regression Results: Analysis of Helpfulness Across Countries – Dependent Variable: Helpful Score

	US		UK		Germany		France		Japan	
	Coefficient	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**
Intercept	.442	.022	.416	.044	.339	.044	.362	.073	1.076	.018
Rating	.030	.016	.082**	.031	.059	.031	.130**	.048	-.263***	.015
Rating Squared	.003	.003	-.006	.005	-.002	.005	-.015*	.007	.039***	.002
Adjusted Character Count	.000***	.000	.000***	.000	.000***	.000	.000***	.000	.000***	.000
Total Votes	.000	.000	.000*	.000	.000*	.000	.003***	.001	.000***	.000
N	11,326		2,978		2,755		1,130		5,143	
R-Squared	.224		.210		.284		.208		.275	

***p<.001
**p<.01
*p<.05

First, to examine the effect of uncertainty avoidance predicted in hypothesis 4, the results for the model applied to the Japanese reviews need to be examined. The model with Japanese reviews is shown to have a negative coefficient for the rating variable ($p<.001$), but a positive coefficient for the squared rating term ($p<.001$). This indicates that the Japanese reviews were found to be more helpful at the extreme ends of the scale (one and five stars) than reviews in the middle of the star rating scale (two-four stars). This result supports hypothesis four and suggests that countries with high levels of uncertainty avoidance, like Japan, find reviews with extreme valence more helpful. When compared to the results of the western countries, Japan is the only country that exhibits this dynamic, suggesting that this effect may be unique to high uncertainty avoidance counties like Japan. This has meaningful consequences for managers and researchers seeking to understand review helpfulness in countries with high uncertainty avoidance. Since the results suggest that these consumers find one and five star reviews to be more helpful retailers need to take this into consideration when displaying reviews on their sites. Making one and five star reviews more prominent in countries with high uncertainty avoidance may improve the shopping experience for consumers. It is also interesting that Japanese consumers on average produce lower mean star ratings compared to their western counterparts as established in the testing of hypothesis 3, but yet actually find five star reviews to be helpful.

The results of these models also show some surprising results regarding the western countries. Prior research concerning online reviews in the US suggests that consumers have a preference for more negative reviews (Sen and Lerman 2007;

Mudambi and Schuff 2010). In looking at the model for US reviews, both the rating and rating squared variables are insignificant. In the UK model, the rating coefficient is actually positive and significant ($p<.01$) while the rating squared coefficient is negative, but insignificant. This suggests that UK customers may actually find positive reviews to be more helpful. These findings run counter to what would be expected to be found based on prior research which may call into question prior assumption regarding valence and helpfulness and provide an opportunity for further research. The effect of adjusted character count was to be significant ($p<.001$) in all countries suggesting that, in line with prior research, longer reviews are found to be more helpful, though the size of the effect is minimal.

Overall, testing a model of review helpfulness across countries and languages has provided new insights to the body of literature on online reviews. This research shows that prior research analyzing review helpfulness may not be applicable across different countries, cultures, and languages. Finding that Japanese reviews have a different helpfulness pattern across star ratings compared to their western counterparts shows that prior assumptions about drivers of review helpfulness may not be applicable when viewed in an international context. Additionally, the finding that star rating did not have a significant effect on helpfulness for US reviews, and a positive effect for UK reviews calls into question the findings from prior research that suggests western consumers have a preference for more negative reviews.

Discussion and Conclusions

The hypotheses and results are summarized in Table 16. The results provide evidence that cultural bias affects measurable and visible aspects of consumer behavior on online review sites. The communication culture of a country may influence the relative verbosity of reviews in that country. Raw character counts showed that Japan, which is associated with a high context communication culture, had the shortest reviews. However, when taking language structure into account, this effect did not hold. This finding shows that both language and culture need to be taken into account when assessing differences in online review volume across countries. In addition to systemic differences in WOM volume across countries, star ratings and helpful votes vary systematically across country sites. More specifically, Japan, which is associated with a weak extreme response bias and strong negative response bias, was shown to have reviews with a lower average star rating when compared to western reviews. This suggests that retailers need to be cautious when comparing star ratings across countries as lower star ratings may be an artifact of cultural influence and not consumer satisfaction. Acquiescence bias, collectivism and uncertainty avoidance were also shown to be influential factors determining the helpfulness of reviews. Japan, which is associated with acquiescence bias and a collectivist culture, showed a higher overall level of review helpfulness. This suggests that culture needs to be considered when looking at helpfulness as a measure of review quality. Uncertainty avoidance was also shown to be a factor in review helpfulness when taking valence into account. Japan, which is associated with a high level of uncertainty avoidance was shown to have a higher level of review helpfulness for extreme reviews. This shows that while Japanese reviews have a higher overall level of helpfulness, valence also needs to be considered when comparing review helpfulness across countries.

Table 16. Chapter 4 Summary of Hypotheses and Results

	Description	Result
H1	The more a language is associated with a high-context culture, the lower the verbosity of online reviews in that language.	**Mixed Support**
H2	Cultural response biases affect online review valence, reflected in the star ratings in that culture's reviews	**Supported**
H3	Online reviews in countries associated with acquiescence bias will have a higher percentage of reviews voted helpful than reviews in other countries.	**Supported**
H4	Extreme reviews will be considered more helpful than moderate reviews in countries associated with high uncertainty avoidance.	**Supported**

Firms should be hesitant to compare the length of review, star ratings, and helpfulness of reviews for a particular product without taking into account differences attributable to language and culture. The results are especially relevant for multinational retailers expanding in Asia. If ratings on Asian language sites appear lower than on non-Asian sites, this might not indicate lower satisfaction with product quality, but rather a cultural response bias. In addition, the findings imply differences in helpful votes across country sites should also be interpreted cautiously. Helpful votes generally indicate the level of quality of a review (Mudambi and Schuff 2010), but the overall quality of the body of reviews cannot simply be compared across country sites without considering that consumer agreement tendency, and level of uncertainty avoidance varies across countries.

Retailers seeking to assess review quality need to utilize multiple measures of quality in cross-country comparisons.

This study develops a theory-grounded way to evaluate the effect of culture and language on verbosity, valence, and helpfulness of online reviews. This approach involved disentangling language and cultural effects, and adding insights on cultural response bias in online consumer behavior and interactions in the context of online reviews. As commerce is increasingly transacted online, and as consumer-to-consumer interactions shape the image of brands and retailers, business strategy depends on a clear understanding of online consumer behavior and social cognition. The findings identify cross-border differences and similarities in consumer behavior regarding online reviews, and develop an enhanced perspective on language and sociocultural norms.

While raw character counts indicated Japanese (high-context) reviews are shorter than those of Western (low-context) countries, the adjusted character counts tell a more nuanced story. The results suggest, that on the surface, the relative verbosity of online reviews across countries falls in line with what could be expected using the high/low context communication framework. When taking language structure into account however, the high/low context communication cultural perspective does not fully account for differences in volume of WOM. This finding can lead to a rethinking of the theoretical relationship between high and low-context cultures and verbosity and provides an opportunity for future research.

Since many global consumer products have star ratings on multiple country sites, this raises a question of the comparability of consumer sentiment across countries. Building on past research on cultural response bias, significant differences in the average star rating of products across countries were found. Reviews in Japan exhibited a significantly lower star rating than reviews in the US, Germany, France, and the UK. This finding cautions managers against making conclusions at the product level. A specific product with a lower star rating in Japan may not be cause for alarm. If the star rating of a product in the UK is higher than in Germany, this does not necessarily indicate UK consumers are more satisfied.

The research has also generated insights on interpreting review quality. Review sites encourage consumers to post reviews that will be perceived as helpful by others, and firms can compare average helpfulness of reviews across countries as a metric of review quality. However, this can be misleading, due to cultural differences in acquiescence. Although Japanese reviews have *lower* average star ratings, they have a much *higher* reported average helpfulness. This may be a new indication of what psychologists have called the dialectical nature of the East Asian self-concept, a self-concept that embraces contradiction, change and a holistic perspective (Spencer-Rodgers et al. 2004; Spencer-Rodgers et al. 2010). Additionally, while Japanese reviews have lower average star ratings and higher reported average helpfulness, the valence of reviews found helpful follow a unique pattern. Japanese reviews were found to be viewed as more helpful when falling on an extreme point on the 1-5 star scale. This provides further evidence that it can be problematic to compare the average helpfulness of reviews at a product level across the country sites. Additional measures of review quality could add value at the product level, to avoid reliance on the single measure of "is this review helpful."

While the findings of this research add insight to potential issues in comparing online review data across countries, the study does have some limitations. The dataset used to test the hypotheses was collected from a single multinational retailer, Amazon.com. It is possible that consumer reviewing behavior may differ across other retail platforms. Future research could address this issue by examining additional data from other retailers. Since Amazon is a US based retailer, it may also be beneficial to examine any potential differences between Amazon and a multinational retailer based outside the US like Taobao. Additionally, the data in this study is limited to five countries (United States, United Kingdom, France, Germany, and Japan). While these selections likely represent meaningful online shopping cultures, this research could be extended to include a more diverse set of country data. Applying this research to retail data from large developing markets like China and Russia may yield additional insights. Additionally, the influence of product type was not examined in this study. Prior research has shown that product type may be an influential factor in online reviewing behavior (Sen and Lerman 2007, Mudambi and Schuff 2010). Future research should examine product type as a potential influence on reviewing behavior in an international context.

Future research can build on these findings to develop finer-grained explanations of differences in consumer culture and behavior. One opportunity for future research is an exploration of online WOM in the context of culture and consumer complaining behavior. Past work on WOM in private settings has shown that collectivist cultures are more likely to spread negative WOM (Chan and Wan 2008), and are more likely to share dissatisfaction with their in-groups (Liu and McClure 2001). The finding that Japanese reviews tend to have a more negative valence than their Western counterparts could support this, but online reviews are neither private nor shared solely with a consumers' in-group. A negative online review could be viewed through Singh's (1988) taxonomy of consumer complaining behavior as voice response (complaining directly to the retailer), private response (sharing information with the online community), or a third party response (complaining to a customer review forum), so additional research is called for. The potential connection between review length and complaining behavior is also worthy of further study, as prior research is inconclusive.

To examine text comments across languages, a multi-method approach of analyzing text comments through in-depth qualitative analysis and machine coding could provide insights into the emotional, cognitive, and structural components of online reviews. In particular, US reviews have more depth of discussion than in other countries, but it is not clear if this is best attributed to the higher level of US experience with online expression, stronger ties to Amazon, incentives for top reviewers, or simply a higher level of consumer interest and engagement. Other market or environmental factors may play a role if accessibility to online reviews is limited. Language and cultural differences are likely to become even more important as online commerce expands to new and emerging markets.

As user-generated content becomes a more influential component of the global consumer's decision making progress, multinational firms need to take a more systematic and strategic approach to analyzing reviews and social media content. Taking language and cultural differences into account is an essential component of listening in to consumers. This study contributes to a theoretical understanding of cross-language differences in consumer behavior regarding online reviews, and provides findings that are statistically and practically significant to international marketing strategy.

CHAPTER 5: CONSUMER IMAGES IN DIGITAL DECISION AIDS: THE SELFIE EFFECT

Introduction

Advancing technology has increased the ability of marketers and consumers to provide and examine product information in a virtual setting. The ease with which consumers can view and capture images is changing the way shoppers view and evaluate products, and themselves. The powerful cameras contained in smartphones that many consumers don't leave home without has dramatically increased the practicality for consumers to take pictures of products, friends, and themselves. The sheer number of consumers with smartphones is overwhelming. In 2014, in the U.S. alone, over 170 million people owned smartphones, including 85% of Millennials (Nielson 2014). The prominence of this technology is now allowing the use of digital decision aids that are more mobile, sharable, and more seamlessly integrate with ordinary shopping behavior.

A visit to a popular review site like Yelp! can illustrate how consumers are utilizing the picture taking ability of their mobile devices in their consumption experience. Imagery has become prominent in online reviews as a consumer restaurant review often contains pictures of their food, the environment, and themselves in addition to a text evaluation and star rating that are traditionally found in online reviews. This practice is particularly common when it comes to online reviews in fashion retailing. An inspection of consumer reviews on popular fashion rental site, renttherunway.com, shows that consumers frequently include pictures of themselves in the dress they rented in their review. Beyond online reviews, newer digital decision aids have heightened the importance of consumer images in the shopping experience. Virtual Try-on is one type of technology that allows consumers to create virtual models of themselves to "try-on" apparel when shopping online (Salfino 2014). One manifestation of this, called virtual mirrors, serves as an example. Using personal electronic devices, virtual mirrors enable consumers to view themselves using or trying on a product without physically doing so in person. One example currently being employed by marketers is L'Oreal's recently developed "Makeup Genius," an online and mobile software application that allows consumers to virtually try on cosmetics using the cameras on their phones or tablets (Stout 2014). Technologies like these illustrate the importance for marketers and researchers to better understand how consumers use and perceive images that may influence product evaluation and purchase decisions.

The way many consumers capture and view images of themselves has changed. Traditionally, the only way people saw their own image was by looking in the mirror or by seeing a picture of themselves that was likely taken by someone else and at some time in the past. It has long been shown that individuals prefer pictures of themselves when they corresponds to their *mirror* image (i.e., what they see when they look in the mirror) rather than their *true* image (i.e., how others see them) because they are more frequently exposed to their mirror image than their true image (Mita et al. 1977).The dramatic changes in available technology and associated consumer usage habits call into question this long standing assumption. Due to the widespread use of smartphones, social media, and the prominence of digital self portraits, commonly known as "selfies" (a photograph that one has taken of oneself), the nature of how people view themselves is changing.

Technology has dramatically increased the frequency and manner in which we see ourselves day to day. The proliferation of smartphones, social media, selfies, and photo sharing has changed the way consumers have become used to seeing themselves. In addition to added opportunity for people to see their true image through pictures that are taken and shared by others, most smartphones default to saving and displaying a true image when a selfie is taken. A consumer who frequently takes digital self portraits likely both sees their image more often, and is more frequently exposed to their true image. This phenomenon has consequences for how consumers view themselves and products in images. Despite this change, academic research has not yet explored how this issue can impact consumer product evaluation and the consequences for digital decision aids.

This research seeks to answer the question: how does a consumer's exposure to their own image influence their evaluation of products and self-images in digital decision aids? The selfie phenomenon and its effect on how consumers evaluate images likely play a key role. Through three experiments, this research examines the influence of image type, and consumer selfie taking on image and product evaluation. The findings make an important contribution to the literature that expands our understanding of the aspects of consumer images that can influence their use in digital decision aids, and opens the door for more research on the subject. Broadly, the results demonstrate that consumer exposure to their own image (through selfie taking) will influence how consumers evaluate products and pictures. More specifically, this research identifies a "selfie effect" that influences consumer opinions of products in images. This effect shows two key aspects. First, consumers who frequently take digital self portraits are shown to evaluate products seen in images more positively. Second, consumers who frequently take selfies prefer products when they are viewed in their true (vs. mirror) image. This research also calls into question the long standing assumption that consumers prefer their mirror (vs. true) image. These findings have significant implications for consumer use and interactions with digital decision aids. To better understand the context of these studies and findings, a brief review of consumer image taking and the digital decision aids impacted is needed.

The ease with which consumers can capture digital images of themselves via selfies facilitates their ability to try on products with more realistic images in a wider variety of environments. Many online retailers are introducing applications that allow consumers to see what their product looks like on them without ever entering a store. Virtual try-on helps internet retailers overcome the limitation of buyer's inability to try on a product before purchase allowing them to better compete with brick and mortar retailers (Varadarajan et al. 2010). Kim and Forsythe (2008) show that virtual try-on technologies can also increase the entertainment value of the online shopping experience. These virtual try-on technologies allow consumers to view products ranging from glasses to make-up on their own picture. Ray-Ban for example offers an iPhone application which uses images taken from a consumer's in-phone camera to create an interactive image for use in the application. With this image consumers can "try-on" different frames to evaluate how they look. The application also allows for shoppers to make purchases and even share the images with friends via social media. Despite the development of new shopping tools like this, there has been limited academic research concerning implications of consumer images used for this technology. What type of images should be used?

Researchers have begun to apply the long standing assumptions about consumer image evaluation to modern technologies. Cho and Schwarz (2010) examine image type in the context of virtual mirrors and found that consumers preferred accessories (such as earrings) and were more likely to recommend them when viewed on the true image of a familiar other. Results are less intuitive however when investigating one's own image. Cho and Schwarz (2012) found that while people liked their mirror image more than their true image, products shown in their mirror image did not enjoy a significant evaluation advantage. It is not entirely clear why the true image vs. mirror image format affected product evaluation on familiar others, but not on the self. Cho and Schwarz (2012) speculate that the frequency with which we see ourselves in pictures and videos may be an explanation, but further research is needed to explore this issue.

Conceptual Development and Hypotheses

To maximize the potential of new technological shopping tools, it is important to understand what influences consumers' perception of themselves while using them. At the same time, the influence of how modern consumers are regularly viewing themselves also needs to be considered. The mere exposure effect refers to the notion that individuals like objects more if they feel more familiar, and that familiarity is a function of the number of prior exposures (whether supra- or subliminal in nature) (Zajonc 1968). In the case of self-images, the relative frequency with which we see ourselves in the mirror increases the ease with which these images are processed (i.e., fluency) compared to less-often-seen true images, which in turn creates feelings of familiarity and liking (Mita et al.

1977). The assumptions of the mere exposure effect, processing fluency, and their effect on consumer image evaluation all need to be considered when evaluating consumer product evaluation in images and the implications for digital decision aids.

It is widely believed that processing fluency, or the ease with which information is processed, creates feelings of familiarity, which translates to enhanced evaluations and increased aesthetic pleasure (Reber et al. 2004; Reber et al. 1998). The easier it is to recall the positive features of a product, the more favorably consumers will evaluate it (Menon and Raghubir 2003). People make judgments about something based on the level of difficulty they have at the time of evaluation and misattribute the fluency of the process to their evaluation of the stimuli (Wilcox and Song 2011). Simply put, people view an object more favorably when processing is fluent (Janiszewski 1993), and prior exposure to a stimulus increases the ease of processing (Labroo et al. 2008).

Processing fluency has been demonstrated to be influential in a number of contexts. Consumers believe that there is a larger price difference between products if the difference is easier to compute (Thomas and Morwitz 2009). Lee and Labroo (2004) show that increased advertising exposures can lead to consumer recognizing a brand more easily. Enhanced processing fluency of a brand can also lead to favorable evaluations and increased brand choice (Lee 2002). For example, consumers have more favorable attitudes toward a product after being exposed to an advertisement for that product (Lee and Labroo 2004).

More generally, "…aesthetic pleasure is a function of processing fluency: any variable that increases the fluency with which an object can be processed also increases the perceiver's aesthetic pleasure (for reviews see Reber et al. 2004; Schwarz 2004)."

(Cho and Schwarz 2010, p 472). The fluency increasing variables of interest in this research are product proximity and selfie frequency. Consumers should be very familiar with seeing themselves in day to day accessories and apparel (hat, earrings, make-up, etc). This familiarity makes it easier for consumers to process images of themselves in such items. Fluency is generated by the ease of generating thoughts and the ease of processing external stimuli (Novemsky et al. 2007). Consumers evaluate products that are easier to process visually more favorably because they mistake the ease of processing for how much they actually like the product (Wilcox and Song 2011) Since consumers are familiar with seeing themselves in common products, it will be easy for them to process their image when they try them on in a shopping situation. Even though they may not be familiar with the exact product they are trying on, their familiarity with their own image with products on should lead to processing ease. Conversely, consumers are not as familiar with seeing products placed on other displays. For example, viewing a product on the blank face or body of a mannequin is much more unfamiliar for consumers than seeing a product on themselves.

This is especially true in today's environment. The widespread use of social media and camera phones has dramatically increased the frequency in which we see ourselves, and see ourselves wearing products. Stylistic properties of images can serve as a source of persuasion (Yang, Zhang and Peracchio 2010), It has been estimated that over half of the population in the United Kingdom have taken selfies and that 35 million are created every month (AdWeek 2015). Since individuals like objects more if they feel more familiar and familiarity is gained by prior exposures (Zajonc 1968), it is likely that consumers who frequently take digital self portraits are even more familiar with their appearance in products and therefore are more likely to positively evaluate products on their person. Conversely, consumers are less likely see products that are in close proximity to them, but not on them, and will be more likely to negatively evaluate them. While a consumer can easily process their own image, or themselves with a product on, when a product is placed near a consumer, it will interrupt the processing fluency of the image they see due to the less familiar arrangement of the image. Seeing a product on the blank face of a mannequin for example, would be more unfamiliar and thus a more difficult image to process than one containing just a product and their own face. The practice of taking selfies is likely to amplify this effect. Selfies have become a phenomenon that has drawn nationwide attention and become a regular practice for many consumers. Frequently capturing digital self portraits increases the frequency of a consumers' exposure to their own image, thereby making it even more familiar to those who do not regularly engage in this practice. Because of the processing ease of viewing a product on their person, which is increased for frequent selfie takers, and the processing interruption that will be caused by a product being nearby, but not on, it is hypothesized that:

H1: When viewing an image of self and product, consumers will like the product more when it is on their person than when it is displayed next to them (e.g., hat next to head versus on head), especially true for consumers who frequently take selfies.

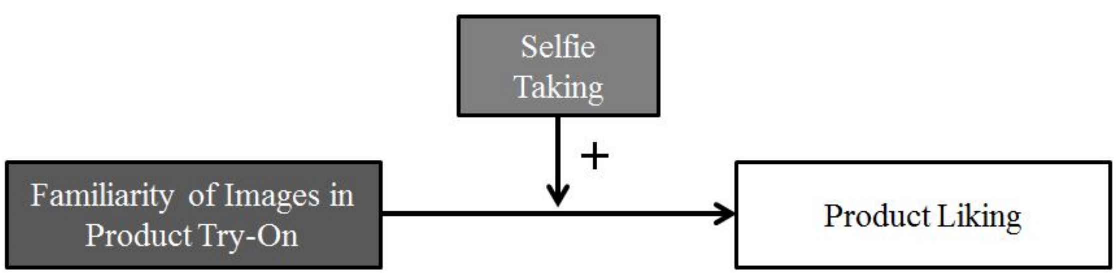
Familiarity leads to attraction (Reis et al. 2011). It is believed that people prefer pictures of themselves when they correspond to their mirror image rather than their true image because they are more frequently exposed to their mirror image than their true image (Mita et al. 1977). It has also been shown that the reverse is true when considering images of a close friend – because we are more frequently exposed to friends' true images and therefore prefer them to friends' mirror images (Mita et al. 1977). Both of these effects can be explained by a fluency account as is demonstrated in the mere exposure effect (Zajonc 1968). It is a fluency difference that is credited for people's preference for mirror self-images to true self-images (Mita et al. 1977). While past research identifies this advantage for mirror images, it is not fully understood why this preference does not translate to evaluations of products on one's mirror (versus true) image (Cho and Schwarz 2012).

The current environment is changing how we see ourselves. At one time, we most commonly saw ourselves when we looked in the mirror (our mirror image). On occasion, we saw pictures of ourselves (true image) that were taken at some point in the past by someone else. Given this dynamic, it is intuitive that people would be more familiar with, process more easily, and prefer our mirror image to our true image. Today however, this dynamic no longer holds. Every time we log into Facebook, snap a selfie, or scroll through recent memories in the gallery of our phone, we are exposed to our true image. With the convenience of mobile digital photography, the frequency in which we see ourselves in pictures (true image) has dramatically increased. It is likely that many people see pictures of themselves more often than they see themselves in the mirror. Due to this phenomenon there has likely been a shift in consumer image preference. Our new greater familiarity with our true image should make it easier for us to process and thus prefer it to our mirror image which we now see less frequently in comparison. As a result, our preference for products in images has also likely changed. Since we are now used to seeing ourselves wearing products in our true image, when trying on a product, consumers will likely prefer the product in their more familiar true image. On most smartphones, selfies are recorded and saved as true images. Due to this design, people who take these images frequently are even more likely to have familiarity and fluent processing with their true image due to the additional exposure to it through selfie taking. Because of this shift in consumer perception and the ability of selfie taking to amplify this effect, it is hypothesized that:

H2: When viewing an image of themselves trying on a product, consumers will like the product more in their true image than their mirror image, especially true for consumers who frequently take selfies.

A general illustration of the conceptual premesis for the hypothesized relationships is provided in Figure 4.

Figure 4: **A Conceptual Model of Image Familiarity and Product Liking**



Study One

Consumer evaluations can be influenced by difficulty experienced when they are making the evaluation (Wilcox and Sun 2011). For example, Menon and Raghubir (2003) show that consumer judgments of a product were more favorable when they only had to remember two product attributes as opposed to a more difficult eight attributes. Since consumers frequently see their own image through pictures taken by others, selfies, and looking in the mirror, they are familiar with seeing various types of products on them in images, making it easy for them to process and therefore evaluate these products. It is likely that consumers who frequently take selfies are even more familiar with seeing themselves in products. This increased familiarity with their own image with products on will increase the processing fluency of these images and strengthen their preference for products when seeing them on themselves. However trying on a product is not the only way people can view products. Consumers are less familiar with images in which there is a product in the image, but not on them, which will result in an less fluent image and less favorable product evaluation. The first study seeks to test hypotheses one and establish that consumers, especially frequent selfie takers, will evaluate a product more favorably when it is viewed on them as opposed to near them. It is expected that consumers will have an easier experience evaluating a product that is on them than next to them on a blank mannequin face leading to higher product evaluations when they are wearing it. This will be especially true for people who frequently take selfies. Additionally, a memory task should make it more difficult to evaluate a product and result in lower product evaluations.

Method

118 undergraduate students (86 male) participated in the study in exchange for course credit. They were randomly assigned to one of the four experimental conditions of a 2 (product-on vs. product-near) by 2 (long number vs. short number) between subjects design. Participants were told that the study was interested in how consumers evaluate products in images. First, participants were given an instruction sheet explaining the procedure and were instructed to remember either a two digit or seven digit number throughout the duration of the study. Remembering the number was designed to make the product evaluation more difficult for participants in the seven digit condition which would then interfere with the participants' ability to fluently process the images they were viewing (Miller 1956). Participants were then instructed to sit at a station with only a mirror in front of them. Participants in the product-on condition found a hat at the station and were asked to try it on and leave it on through the duration of the study. Participants in the product-near condition found a hat placed on a mannequin head facing the mirror at the station and were instructed not to touch or move the hat and mannequin throughout the study. In this condition, the hat and mannequin were placed so that they would appear just to the side of the participant at approximately shoulder height in the reflection in the mirror. A blank wall was behind the participants to ensure that they and the hat/mannequin were the only images in the reflection. A plain winter hat in the University's colors was used to minimize potential product related effects and increase familiarity with the features of the product.

After being assigned to the conditions and seated with the product on or near them, participants were asked to respond to a questionnaire while referring to the image in the mirror in front of them. While looking at the reflection participants responded to several questions concerning their opinion of the hat on a seven point scale. Questions included: how much do you like the hat (1=not at all 7=very much), how attractive do you think the hat is (1=not at all attractive 7=very attractive), how likely would you be to purchase the hat (1=very unlikely 7=very likely), how likely would you be to recommend the hat to a friend (1=very unlikely 7=very likely), how good do you look in the hat (1 not good at all 7=very good). All five of these items were averaged to create a product appreciation index (α = .88). Participants were also asked how many selfies they typically take in a week and for demographic information. Finally, participants were asked what the number they were instructed to remember was and thanked for their participation in the study.

Results and Discussion

After reviewing participant responses, three participants were found to have not fully completed the questionnaire and were removed from the analysis. Initial testing revealed that the long number versus short number manipulation and its interaction with the hat on/hat near manipulation did not have a significant effect. Accordingly, the results from participants were collapsed across the cognitive load manipulation for analysis. To test hypothesis one, a regression was conducted using the five-item product appreciation index as the dependent variable. Hat on (0=hat near, 1=hat on), the number of selfies a participant takes in a week (mean centered), and their interaction term served as predictors of how much participants would like the hat in the reflection. Hat on was found to have a significant main effect (B = .492, t(3, 111 = 2.14), $p<.05$) indicating that consumers prefer seeing themselves wearing the hat as opposed to seeing it on the mannequin, providing partial support for hypothesis one. Number of selfies was also shown to have a significant main effect (B = .024, t(3, 111 = 2.00), $p<.05$) indicating that the more selfies a participant takes the greater their appreciation of the hat. The interaction term however was not significant (t(3, 111) = -1.01) in the analysis. The regression results are displayed in Table 17. Despite the lack of support for the hypothesized interaction between selfie taking and the hat on/hat near condition, it is interesting to establish that consumer selfie taking habits may influence their evaluation of products in images. It is possible this result indicates that consumers who take a lot of digital self portraits may be more used to viewing their own image and therefore more positively evaluate products shown in an image with them regardless of where the product is placed in the image because they are used to and more comfortable viewing images of themselves. Studies two and three seek to further explore this possible “selfie effect”.

Table 17. Regression Results: Chapter 5 – Study 1
Dependent Variable: Product Appreciation Index

Variable	Coefficient	SE	t-value	Sig.
Hat On	0.492	0.230	2.139	0.035
Number of Selfies	0.024	0.012	2.001	0.048
Hat On*Number of Selfies	-0.014	0.014	-1.013	0.313
N = 115				
r^2 = .09				

Study Two

Study two seeks to further investigate the finding of the main effect of selfie taking in study one. Consumers regularly view products online, both from retailers and in online reviews, and it is important to understand factors that may influence how consumers evaluate products in these images. Study one demonstrated that consumers who take a lot of selfies liked a product more regardless of whether the product was on them or near them. Study two seeks to further evaluate this relationship by examining if consumers who take a lot of selfies will also more highly evaluate a product, regardless of image type, when viewing the product online. This removes the image of the consumer from the equation and will help verify if consumers who take a lot of selfies have will more positively evaluate a product even when their own image is not present.While this study does not directly test the proposed hypotheses, it is exploratory in nature and will help clarify the selfie effect found in study one. Additionally, if confirmed, the results may add an interesting contradiction to prior literature. It has previously been found that online images of products containing human images improve consumer attituted towards products and online shopping sites (Wang et al. 2014). Human images online have also been shown to enhace image appeal and increase online trust (Cyr et al. 2009). Study two will help to determine if consumer selfie taking habits influence this relationship and possibly create opportunities for future research.

Method

156 undergraduate students (77 male) participated in the study in exchange for course credit. The study was conducted through an online survey to better align the image evaluation with an online shopping scenario and to remove the participants own image from the experiment. Participants were randomly shown a scarf either alone, in a plain image as it might be displayed by an online retailer, or worn by another consumer. A scarf in the University's colors was used to minimize potential product related effects and increase familiarity with the features of the product. The image of the scarf on another consumer was taken of a college age (female) volunteer, from the abdomen up, with a plain background to avoid interference from any other objects in the picture. While referring to the random image they were shown, participants responded to several questions concerning their opinion of the scarf on a nine point scale. Questions included: how much do you like the scarf (1=not at all 9=very much), how attractive do you think the scarf is (1=not at all attractive 9=very attractive), how likely would you be to purchase the scarf (1=very unlikely 9=very likely), how likely would you be to recommend the scarf to a friend (1=very unlikely 9=very likely), how good do you look in the scarf (1 not good at all =very good). All five of these items were averaged to create a product appreciation index (α = .91). Participants were also asked how frequently they take selfies on a nine point scale (1=very infrequently, 9=very frequently). At the end of the survey participants were asked what color the scarf in the image was to ensure they had viewed and evaluated the image in the study. After reviewing the responses, three

participants were found to have responded to the survey incorrectly and were removed from the analysis.

In a procedure similar to that in study one, the responses to questions concerning how appealing the hat was to the participants were averaged to create a measure of how much they liked the scarf in the reflection which will be referred to as the product appreciation index. To build upon the findings of study one, a regression was conducted with product appreciation index as the dependent variable. Image type (0 = product alone, 1 = product on a person), selfie frequency (continuously measured and mean centered) and their interaction term were included as predictors of how much participants like the scarf.

Results and Discussion

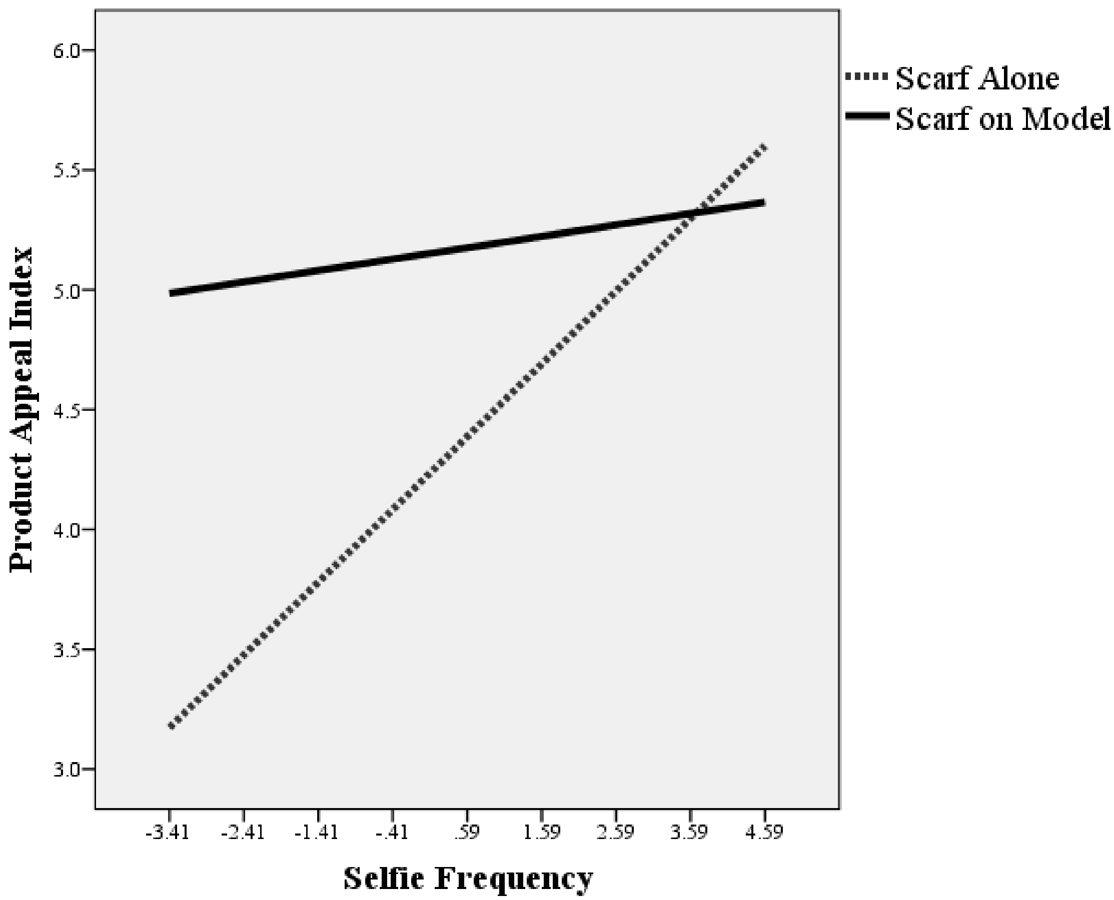
Image type was shown to have a significant main effect (B = .931, t(3, 149 = 3.22), $p<.01$). This result indicates that consumers evaluated the product more favorably when it was shown on another consumer rather than by itself. Selfie frequency was also shown to have a significant main effect (B = .303, t(3, 149 = 3.28), $p=.001$). This indicates that people who more frequently take selfies evaluated the product more favorably. These results are qualified by an interaction of image type and selfie frequency (B = -256, t(3, 149 = -2.01), $p<.05$). The regression results are displayed in Table 18. An inspection of the interaction effect shows that low selfie takers liked the product more when on model than when alone, while high selfie takers liked scarf same regardless of whether alone or on other person.

Table 18. Regression Results: Chapter 5 – Study 2
Dependent Variable: Product Appreciation Index

Variable	Coefficient	SE	t-value	Sig.
Image Type	0.931	0.289	3.219	0.002
Selfie Frequency	0.303	0.093	3.276	0.001
Image Type*Selfie Frequency	-0.256	0.122	-2.099	0.037
N =155				
r^2 = .12				

The interaction shows that while participants who frequently take digital self portraits evaluate the product favorably regardless of image type, participants who viewed the product alone had a much less favorable evaluation of the product if they infrequently take selfies. Refer to Figure 5 for an illustration of this effect.While further research will be needed to explore this effect, study two confirms the surprising “selfie effect” found in study one that demonstrates consumers who frequently take selfies have higher product evaluations regardless of image type. This result is particularly interesting as prior research has suggested that consumers prefer human images when shopping online (Cyr et al. 2009; Wang et al. 2014). The results of study two indicate that consumer selfie taking may add an additional variable to consider in this assumption. In conjunction with study one, this study also suggests that consumers prefer products when then are displayed on a person than displayed in other formats.

Figure 5. Study 2 Interaction Effect of Selfie Frequency and Image Type



Study Three

Studies one and two establish that consumers evaluate products more positively when they are viewed on a person (whether it be themselves or another). This effect is especially true for consumers who do not frequently take selfies. It was also found that consumers who frequently take selfies generally evaluate products more positively than those who do not. Study three is a two part experiment that seeks to extend these findings by examining consumer evaluations of products in different image types (mirror vs. true) and by different picture takers (self, other). Study three consists of a between subjects experiment and a follow-up within subjects experiment. Both parts of the study will test hypothesis two and help to establish preferred image types and perspectives through two complimentary designs.

Between Subjects Experiment

The between subjects experiment will test hypothesis two and help to establish preferred image types and perspectives for consumer product evaluations. Given the frequency of which people now see themselves in pictures, it is expected that consumers will more positively evaluate products when viewed in their true (vs. mirror) image. It is also expected that this effect will be strengthened for consumers who frequently take selfies since selfies provide for frequent exposure to their true image, making them more fluent.

Method – Between Subjects

The first part of study three utilizes a between-subject design. For the study three, 102 undergraduate students (63 male) participated in the study in exchange for course credit. They were randomly assigned to one of the four experimental conditions of a 2(mirror, true) by 2(selfie, other) between subjects design. Participants were told that the study was interested in determining how consumers evaluate products in images. First, participants were informed that they would be photographed while trying on a product in the study and that they would be required to participate in the second part of the study, to take place online 10 days after the experiment.

After being explained the procedure, participants were asked to try on the hat that was used in study one. While wearing the hat, a picture of the participant was taken using the camera on a late model smartphone. For participants in the selfie condition, they were asked to take a selfie wearing the hat. The camera phone was preset so that it would capture either a mirror image or true image depending on the assigned condition. Participants in the other condition had their picture taken by the researcher while they were wearing the hat. The camera phone was again preset so that it would capture either a mirror image or a true image depending on the assigned condition. To summarize, a picture of the participant in the hat was captured for one of the four assigned conditions: selfie/mirror, selfie/true, other/mirror, and other/true.

In all conditions the picture was taken in the same location in front of a bare, neutral colored wall with the same lighting to ensure consistent pictures from participant to participant. To ensure the same perspective, pictures taken by the researcher were taken from a position approximately an arm's length from the participant to approximate the same range the images in the selfie condition would be taken. Since prior research has shown that facial expression (smiling vs. neutral) can influence a consumers image preference (Cho and Schwarz 2012), participants in both conditions were asked to choose whichever facial expression they usually prefer in pictures.

After taking the picture with the hat on in each of the conditions, the phone with the image displayed was placed on an otherwise empty desk next to a questionnaire. Participants were asked to complete the questionnaire while referencing the picture. As in the first two studies, the questionnaire asked participants to evaluate the product they tried on through several measures. Questions included: how much do you like the hat (1=not at all 7=very much), how attractive do you think the hat is (1=not at all attractive 7=very attractive), how likely would you be to purchase the hat (1=very unlikely 7=very likely), how likely would you be to recommend the hat to a friend (1=very unlikely 7=very likely), how good do you look in the hat (1= not good at all 7=very good). An average of all five of these measures were used to create a product appeal index (α = .83). Participants were also asked how many selfies they typically take per week. After completing the questionnaire, an additional picture of the participant was taken in the hat by the opposite picture taker condition that they were assigned. Participants were then debriefed, thanked, and told that they would receive an email with a link to the second part of the study in 10 days.

Results and Discussion - Between Subjects Experiment

After reviewing the images, it was found that the pictures of two participants were blurry, which was not noticed upon initial inspection. These participants were removed from the analysis. To test hypothesis two, a regression using the product appreciation index as the dependent variable was conducted. Image type (mirror=0, true=1), picture taker (other=0, self=1), number of selfies taken in a week (continuously measured and mean centered), all two-way interactions, and the three-way interaction were used to predict participants' evaluation of the hat.

The main effects of image type, image taker, and number of selfies were not significant. The two-way interaction between image taker and number of selfies as well as the three-way interaction were also insignificant. The interaction of image type and number of selfies however was a significant predictor (B = .888, t(7, 94 = 2.24), p<.05). The regression results are displayed in Table 19. This finding provides partial support for hypothesis two by illustrating that consumers who take a lot of selfies evaluate products they try on more favorably when they are seen in their true (vs. mirror) image. This is a significant finding as it has previously been assumed that consumers preferred their mirror image. Introducing the selfie effect to the analysis also demonstrates that the selfie taking phenomenon is an important consideration when evaluating how consumers see themselves.

Table 19. Regression Results: Chapter 5 – Study 3 – Between Subjects
Dependent Variable: Product Appreciation Index

Variable	Coefficient	SE	t-value	Sig.
Image Type	0.228	0.387	0.591	0.556
Image Taker	0.150	0.382	0.393	0.696
Number of Selfies	-0.307	0.288	-1.066	0.289
Image Type*Image Taker	0.206	0.549	0.375	0.574
Image Type*Number of Selfies	0.888	0.397	2.240	0.027
Image Taker*Number of Selfies	0.224	0.397	0.564	0.574
Image Type*Image Taker*Number of Selfies	-0.708	0.517	-1.369	0.174

N =101
r^2 = .09

Within Subjects Experiment

The second part of the study seeks to build upon the results of the between subjects experiment by examining the effects of image type and picture type through a within subjects analysis. This experiment is expected to build additional support for hypothesis two by showing that when participants are asked to view images from all four conditions put forth earlier (selfie/true, selfie/mirror, other/true, and other/mirror) participants will evaluate the product more favorably in the true image conditions, especially if they are frequent selfie takers.

Method – Within Subjects

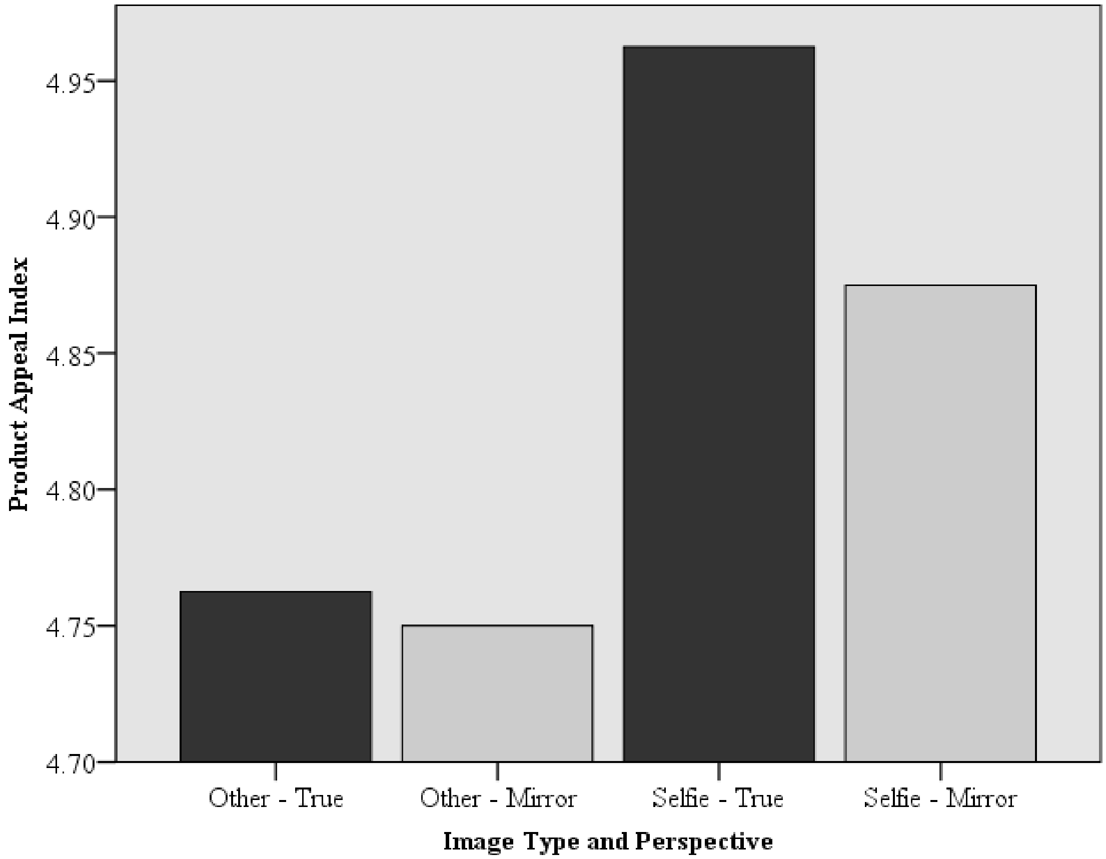
Ten days after participating in the between subjects experiment, participants received an email with a link to complete an additional online survey. In the online survey, participants were presented with, in random order, four images of themselves in the hat used in the earlier study. The images represented all four of the image/pictures types used in the 2x2 within participants design from the first part of the study (mirror/other, mirror/selfie, true/other, true/selfie). While care was taken to produce consistent images in the first part of the study, before the images were placed in the follow up survey, all of the images were cropped to be made the exact same size and contain the same proportions for every image of every participant. For participants originally in the mirror (true) condition, photo editing software was used to flip the picture creating a true (mirror) version of the picture. Participants were asked to individually evaluate the hat in each picture on the same scale used in the first part of the study. After individually evaluating each of the pictures, participants were shown all four pictures at the same time, andasked to rank them in order of preference.

Results and Discussion – Within Subjects

Twenty participants dropped out and did not participate in the follow up study. Two participants were found to have completed the online questionnaire incorrectly and were removed from the analysis, leaving 80 participants included in part two of the study. To further test hypothesis two, a repeated measures ANOVA was performed on the four image conditions with the product appreciation index as a repeated measure. The analysis did not yield a significant result and does not provide support for hypothesis two.

Despite the insignificant result, a review of the distribution of the responses does provide some insight. A visual of the ANOVA results is provided in Figure 6. The directionality of these results is in line with what is predicted by hypothesis two. The selfie/true condition shows to be the most positive product evaluation of the four conditions. While insignificant, the selfie/true condition is the format that selfie takers are likely most familiar with and could be a valuable format to explore further. This observation provokes what could be an interesting area for future research.

Figure 6. Study 3 – Within Subjects Results – Product Appreciation Index by Image Type and Perspective



General Discussion

This research fills a gap in the literature by examining the role of selfie taking and the modern consumers' exposure to their own pictures when considering product evaluation in images. These findings have implications for marketers and provide opportunities for future research. First, this study establishes the "selfie effect". This effect suggests that consumers who frequently take selfies are generally more likely to positively evaluate products seen in multiple types of images. The selfie effect also influences consumer preference for image type. Consumers who frequently take selfies more strongly favor their true image in product evaluation. It has long been assumed that due to increased familiarity, and therefor increased processing fluency, that consumers prefer their mirror image. This research takes a step forward by acknowledging that modern media and technology has changed what consumers are most familiar with and consequently are changing their preference for how they see themselves. This is shown to be particularly true for consumers who frequently take selfies. The effect of the selfie taking phenomenon has not been established in academic literature. While further research on the subject needs to be conducted, this research takes an important step toward identifying the effects of the changing way in which consumers view themselves and its potential effect on marketers and researchers. A summary of the findings from each study are shown in Table 20.

Table 20. Chapter 5 Summary of Results

Study	Key Findings
One	Significant main effects indicate that consumers prefer a product when it is on them versus near them and that consumers who frequently take selfies like a product more than those who don’t.
Two	Significant main effects indicate that consumers evaluate a product more favorably when it is displayed on another consumer, and that consumers who more frequently take selfies like a product more than those who don’t. A significant interaction indicates that while low selfie takers like a product more when shown on a consumer, high selfie takers like the product similarly regardless of whether it is on a consumer or alone.
Three: Between Subjects	A significant interaction indicates that consumers who frequently take selfies evaluate products more favorably when displayed on their true (vs. mirror) image.
Three: Within Subjects	While no significant results were found, the distribution of responses suggests that products displayed as a true (vs. mirror) image in a selfie (vs. other) are evaluated most positively.

These findings also have meaningful consequences for consumer use of digital decision aids. In particular, the findings should prove useful to researchers and marketers concerned with online consumer reviews and virtual try-on technologies. Consumers are increasingly enabled and encouraged to include pictures in the online reviews they write. The results of study two show that people more positively evaluate a product when it is shown on another consumer than when it is displayed by itself. This could suggest that that online shoppers would benefit from online reviews that include pictures of the reviewer demonstrating the product versus reviews that are text alone, or listings that only contain retailer provided product images. Further research is needed to explore this finding specifically in the context of online reviews, but these results suggest that consumer images in reviews may be a fruitful avenue for future exploration. This research also has implications for virtual try-on technologies. Virtual mirrors for example are reliant on consumer images for their function and effectiveness. Understanding that consumers may now have a preference for their true image in product evaluation and that this effect is stronger for consumers who frequently take selfies, is an important consideration in the use and design of these tools. While intuitively, a virtual "mirror" should display a consumers mirror image, according to the findings of this study the tool may be better served if it ensures that products are virtually tried on in a true image.

Despite the useful findings of this research, these studies do have several limitations. This research was conducted using a very limited product selection (a hat and a scarf) in the experiments. While this procedure was chosen to minimize potential impacts of product variety on consumer evaluation, it also limits the generalizability of the findings. Limiting the study pool to undergraduate students also limits the generalizability of the findings. Undergraduates may be among the heaviest users of smart phones and social media which could lead to disproportionate exposure to their own image when compared to other segments of consumers. It is possible that other demographic groups may respond differently to the stimuli used in this research. Perhaps most importantly, there limitations to using selfies as a stimulus in the experiments. Reasonable precautions were taken to ensure that both selfies and pictures taken by the researcher were shot with a consistent angles and proportions in a controlled environment, but it is nearly impossible to ensure that all images are truly consistent among participants. It is possible that technological solutions and photo editing may be able to further mitigate this issue in future research.

The findings and limitations of this research present opportunities rich for future research. The manner and frequency in which consumers are seeing themselves are changing. Academic research has not yet considered the effect selfie taking has not only on product evaluation, but other aspects of the consumption experience as well. Despite the large body of research concerning online consumer reviews, the evaluation of consumer supplied images in reviews presents an under researched area in the literature. The findings here suggest consumers respond favorably to products displayed on other consumers. This concept could be extended to specifically evaluate this effect in the context of online reviews. Understanding how and when consumers may benefit from the inclusion of images in online reviews would add to literature and provide a useful consideration for managers. Perhaps most glaringly, further research needs to be conducted to determine what factors are behind the selfie effect found in this research. While the findings presented here strongly suggest that there are differences between consumers who frequently take selfies and those who do not, it is unclear what the antecedents to selfie taking actually are. Understanding the behavior behind selfie taking is a vital next step to better understanding additional consequences the selfie effect may have on consumer interactions with digital decision aids.

CHAPTER 6: CONCLUSION

The four essays of this dissertation add to the body of marketing literature by furthering our understanding of how consumers interact with digital decision aids. The findings of each chapter both make significant contributions and spur opportunities for future research. The influence that content and context have on consumer interactions with digital decision aids are clearly demonstrated through a diverse yet intertwined set of studies that approach and explain important elements of these interactions from different angles. The essays of this dissertation show that elements of content in digitial decision aids such as: length, valence, helpfulness, language, meaning of text, and the nature of images, are all important considerations when examining consumer interactivity. At the same time elements of digitial decision aid context such as: community norms, website location, national culture, and consumer habits, are also shown to be important considerations in determining consumer interactivity. The contributions made and opportunities created for future research are outlined in this chapter.

Chapter two identifies key dynamics that can influence consumer interactions with online reviews shortly after a product is released. This research furthers understanding of what factors generate interactivity within an online review community. Most notably, Chapter two introduces community norms, and violations of them, as a factor that stimulates consumer action within the review environment. Prior research had identified social dynamics as a factor that can influence behavior within the online review environment. The findings of Chapter two take this a step further by both applying social dynamics to review interactivity and identifying community norms as a key driver behind consumer interactions with online reviews. These contributions are made possible in part by examining reviews in a continuous manner from the time products are released which represents a methodological contribution to online review research as well. In sum, Chapter two has established that aspects of online review content can influence consumer interactivity within the context of a dynamic online review community.

Chapter two also creates and identifies areas conducive to future research. The examination of a reviewer's history in the analysis suggests that a reviewer's propensity to generate highly interactive reviews is a factor that contributes to review interactivity. However, little research has been done to explore the details of reviewer characteristics that enables them to write highly interactive reviews. What is it about these people that drive them to write reviews that draw a response from the community? Research that takes an in-depth look at reviewer characteristics could help to answer this question and further add to the body of research. The identification of community norms as an important factor in the review environment also provokes opportunities for further work. Do different review forums have different sets of norms and are they applied or enforced in different ways? Further research in this area could build upon the findings set forth in Chapter 2. While this research examined the content of reviews and how they influenced interactions, a better understanding of these interactions could be gained by looking at the specific nature of the conversations that are initiated. A deeper examination of the textual content of both reviews and comments could also lead to a better understanding of the specific types of norm violations that generate interactions with reviews.

While Chapter two builds upon work that has been done in traditionally researched review environments, Chapters three and four take the examination of the content of online reviews abroad and examine it in an international context. Research concerning online reviews outside of the United States is very underrepresented in the literature, and this research takes a step toward addressing this gap. The key contribution of Chapter three is the development of the Character Count Language Index (CCLI) which allows for the normalization of written online review content across languages and facilitates more accurate comparisons of user generated content across countries. This index provides the opportunity for researchers to begin to extend current single language based research models using review length as a variable internationally. Given the lack of literature concerning online reviews in an international context, Chapter three provides the opportunity to extend current and future research abroad.

Chapter four begins to utilize the opportunity created in Chapter three by exploring the influence of language and culture on online review content in an international context. The research applies the CCLI developed in Chapter three in an analysis of cultural influence on the relative verbosity of reviews across countries and shows that both language structure and communication context culture may be contributing factors. Findings also indicate that cultural response bias may lead to star ratings of reviews differing significantly between countries, meaning that star ratings may not provide for an equal comparison of customer satisfaction across countries. Cultural differences were also shown to influence how consumers rate the helpfulness of reviews in different countries. These findings provide helpful contributions to both the body of literature and managers while providing opportunities for future research. Chapter four begins to expand online review research internationally and shows that prior models used in single country analyses may be influenced by culture and language, and thus may need to be reconsidered when looking at consumer reviewing behavior in an international context. This research opens the door for researchers to expand prior methods and findings abroad.

While Chapters three and four expand research on digital decision aid content to an international context, star ratings and review text, while highly researched, are not the only forms of content present in digital decision aids. The abundance of smartphones and modern technology allows consumers to take pictures of places, products, and themselves at any time. A quick look at any review site will show that these consumer images frequently end up in online consumer reviews. The use of consumer pictures extends well beyond online reviews however. Modern technology is now allowing the use of digital decision aids that are more mobile, sharable, and seamless with normal shopping behavior, creating different contexts for consumer evaluation. The ease with which consumers can capture digital images of themselves via 'selfies' even facilitates their ability to "try on" products in a wider variety of environments. Chapter five examines the role of selfies and consumers exposure to their own pictures when considering product evaluations in images in the context of digital decision aids.

The findings of Chapter five help to build academic work in the nascent research area of selfie taking and the influence of consumer images in the context of digital decision aids. The study establishes a "selfie effect" that suggests consumers who more frequently take selfies generally evaluate products in images more positively. A second impact of the selfie effect is that consumers who frequently take selfies more strongly prefer their true image when evaluating products. This shows that image preference type can play a role in the effectiveness of the images seen in digital decision aids and should be considered by markers who employ these technologies. This research also takes a step forward by acknowledging that how consumers view themselves in images is changing due to usage of modern personal electronic devices. This finding challenges many previous assumptions about how consumers prefer to view themselves.

Chapter five also presents great opportunity for further research evaluating consumer image use in digital decision aids. Research concerning the inclusion of consumer generated images in online reviews is very limited. A research contribution that could further link the online review literature to the use of consumer images is a needed addition to the body of research. Researchers and marketers could both benefit from understanding how, when, and what type of consumer images would be most beneficial for inclusion in online review content. Additionally, while this researcher establishes a connection between selfie taking and product evaluation in images, it is not clear how consumers respond to seeing selfies of others. Understanding this dynamic could help further explain what images are most beneficial for inclusion in digital decision aids. More broadly, this research suggests that there are differences between consumers who take selfies and those who do so less frequently, but it remains unclear what the antecedents of selfie taking actually are. Exploring this issue further could provide great opportunity for insight on the behavior of modern consumers.

In sum, the four essays of this dissertation make substantial contributions to an important topic in the field of marketing by furthering our understanding of how consumers interact with digital decision aids and provide opportunity for fellow researchers to expand upon this work in the years to come. The essays further understanding of how specific elements of content present in digital decision aids and the context in which they are encountered them drive consumer interactivity with these important tools. As technology continues to evolve, more and more opportunities to examine consumer interaction with new tools will present themselves. Even today, consumers have the ability to walk into a store, have their body scanned to create exact measurements, and get personalized product recommendations based on the analysis. Tools like these present opportunities for current research, and provide a glimpse of what the future may hold for researchers and consumers.

REFERENCES CITED

AdWeek (2015), Surprise statistic of the day: More 'selfies' are shared on Twitter than Instagram. http://www.adweek.com/socialtimes/selfie-statistics/489565 Accessed 5 May 2015.

Aiken, Milam, John Wee, and Mahesh Vanjani (2011), A web-based multilingual meeting system: Breaking the language barrier, *Business Research Yearbook*, 28(1): 71-75.

Agarwal, Ashish, Kartik Hosanagar, and Michael D. Smith (2011), Location, location, location: An analysis of profitability of position in online advertising markets, *Journal of Marketing Research*, 48(4), 1057-1073.

Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood (1997), Interactive home shopping: Consumer retailer, and manufacturer incentives to participate in electronic marketplaces, *Journal of Marketing*, 61(3), 38-53.

Alexa Internet (2013), www.Alexa.com. Accessed 24 July 2013.

Archak, Nikolay, Anindya Ghose, and Panagiotis G. Ipeirotis (2011), Deriving the pricing power of product features by mining consumer reviews, *Management Science,* 57(8), 1485-1509.

Baack, Daniel W. and Nitish Singh (2007), Culture and web communications, *Journal of Business Research,* 60(3), 181-188.

Bagozzi, Richard P., and Utpal M. Dholakia (2002), Intentional social action in virtual communities, *Journal of Interactive Marketing,* 16(2), 2-21.

Baumgartner, Hans and Jan-Benedict EM Steenkamp (2001), Response styles in marketing research: A cross-national investigation, *Journal of Marketing Research,* 38(2), 143-156.

Berger, Johan, Alan T. Sorensen, and Scott J. Rasmussen (2010), Positive effects of negative publicity: When negative reviews increase sales, *Marketing Science,* 29(5), 815-827.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), A theory of fads, fashion, custom, and cultural change as informational cascades, *Journal of Political Economy,* 100(5), 992-1026.

Brannen, Mary Yoko (2004),When Mickey loses face: Recontextualization, semantic fit, and the semiotics of foreignness, *Academy of Management Review,* 29(4), 593-616.

Brown, Jo, Amanda J. Broderick, and Nick Lee (2007), Word of mouth communication within online communities: Conceptualizing the online social network, *Journal of Interactive Marketing,* 21(3), 2-20.

Brynjolfsson, E., Dick, A. A., and Smith, M. D. (2010). A nearly perfect market? *Quantitative Marketing and Economics*, 8(1), 1-33.

Bughin, Jacques, Jonathan Doogan, and Ole Jørgen Vetvik, (2010), A new way to measure word-of-mouth marketing, *McKinsey Quarterly,* 2, 113-116.

Burgmann, Inga, Philip J. Kitchen, and Russell Williams (2006), Does culture matter on the web?, *Marketing Intelligence & Planning,* 24(1), 62-76.

Chabowski, Brian R., Saeed Samiee, and G. Tomas M. Hult, (2013), A bibliometric analysis of the global branding literature and a research agenda, *Journal of International Business Studies,* 44(6), 622-634.

Chan, Haksin, and Lisa C. Wan (2008), Consumer responses to service failures: a resource preference model of cultural influences, *Journal of International Marketing,* 16(1), 72-97.

Chen, Pei-Yu, Shin-yi Wu, and Jungsun Yoon (2004), The impact of online recommendations and consumer feedback on sales, *ICIS 2004 Proceedings,* Charlottesville, Virginia, Paper 58.

Chen, Stephen, Ronald Geluykens, and Chong Ju Choi, (2006), The importance of language in global teams: A linguistic perspective, *Management International Review,* 46(6), 679-696.

Chen, Yi-Fen (2008), Herd behavior in purchasing books online, *Computers in Human Behavior*, 24, 1977-1992.

Chen, Yubo, Qi Wang, and Jinhong Xie (2011), Online social interactions: A natural experiment on word of mouth versus observational learning, *Journal of Marketing Research,* 48(2), 238-254.

Cheng, Yi-Hsiu, and Hui-Yi Ho (2015), Social influence's impact on reader perceptions of online reviews, *Journal of Business Research,* 68(4), 883-887.

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

Chidlow, Agnieszka, Emmanuella Plakoyiannaki, and Catherine Welch (2014), Translation in cross-language international business research: Beyond equivalence, *Journal of International Business Studies,* 45(5), 562-582.

Chiswick, Barry R., and Paul W. Miller (2005), Linguistic distance: A quantitative measure of the distance between English and other languages, *Journal of Multilingual and Multicultural Development,* 26(1), 1-11.

Chintagunta, Pradeep K., Shyam Gopinath, and Sriram Venkataraman (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.

Cho, Chang-Hoan, and Hongsik John Cheon (2005), Cross-cultural comparisons of interactivity on corporate web sites: the United States, the United Kingdom, Japan, and South Korea, *Journal of Advertising,* 34(2), 99-115.

Cho, Hyejeung and Norbert Schwarz (2010), I like those glasses on you, but not in the mirror: Fluency, preference, and virtual mirrors, *Journal of Consumer Psychology*, 20(4), 471-475.

Cho, Hyejeung and Norbert Schwarz (2012), I like your product when I like my photo: Misattribution using interactive virtual mirrors, *Journal of Interactive Marketing*, 26(4), 235-243.

Choi, Incheol, Richard E. Nisbett, and Ara Norenzayan (1999), Causal attribution across cultures: Variation and universality, *Psychological Bulletin,* 125(1), 47.

Cosnier, J., Dols, J.M.F., and Fernandez, A.J. (1986), The verbalization of emotional experiences. In Scherer, Wallbott, Summerfield, (Dir.), *Experiencing emotion: A Crosscultural study*, Cambridge University Press; 117-128, Edition de la Maison des Sciences de l'Homme.

Craswell, Nick, Onno Zoeter, Michael Taylor, and Bill Ramsey (2008), An experimental comparison of click position-bias models, *Proceedings of the International Conference on Web Search and Web Data Mining,* Palo Alto, California, 87-94.

Cyr, Dianne, Milena Head, Hector Larios, and Bing Pan (2009), Exploring human images in website design: a multi-method approach, *MIS Quarterly,* 33(3), 539-566.

Dawar, Niraj, Philip M. Parker, and Lydia J. Price (1996), A cross-cultural study of interpersonal information exchange, *Journal of International Business Studies,* 497-516.

de Almeida, Stefânia Ordovás, Utpal M. Dholakia, José Mauro C. Hernandez, and José Afonso Mazzon (2014), The mixed effects of participant diversity and expressive freedom in online peer-to-peer problem solving communities, *Journal of Interactive Marketing,* 28(3), 196-209.

de Mooij, Marieke, and Geert Hofstede (2011), Cross-cultural consumer behavior: A review of research findings, *Journal of International Consumer Marketing,* 23(3), 181-192.

de Valck, Kristine, Gerrit H. van Bruggen, and Berend Wierenga (2009), Virtual communities: A marketing perspective, *Decision Support Systems*, 47, 185-203.

Dellarocas, Chrysanthos, Xiaoquan Zhang, and Neveen Awad (2007), Exploring the value of online product reviews in forcasting sales: The case of motion pictures, *Journal of Interactive Marketing*, 21(4), 23-45.

Dhar, Vasant, and Elaine A. Chang (2009), Does chatter matter? The impact of user-generated content on music sales, *Journal of Interactive Marketing,* 23(4), 300-307.

Diamantopoulos, Adamantios, and Nicolas Papadopoulos (2010), Assessing the cross-national invariance of formative measures: Guidelines for international business researchers, *Journal of International Business Studies,* 41(2), 360-370.

Dow, Douglas, and Amal Karunaratna (2006), Developing a multidimensional instrument to measure psychic distance stimuli, *Journal of International Business Studies,* 37(5), 578-602.

Duan, Wenjing, Bin Gu, and Andrew B. Whinston (2008), Do online reviews matter? An empirical investigation of panel data, *Decision Support Systems,* 45, 1007-1016.

Dwyer, Paul (2007), Measuring the value of electronic word of mouth and its impact in consumer communities, *Journal of Interactive Marketing,* 21(2), 63-79.

Ethnologue: Languages of the World. http://www.ethnologue.com/statistics/size Accessed 10 October 2014.

Farley, John U., and Donald R. Lehmann (1994), Cross-national "laws" and differences in market response, *Management Science,* 40(1), 111-122.

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

Floyd, Kristopher, Ryan Freling, Saad Alhoqail, Hyun Young Cho, and Traci Freling (2014), How Online Product Reviews Affect Retail Sales: A Meta-analysis, *Journal of Retailing,* 90(2), 217-232.

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

Forsythe, Sandra M., and Bo Shi (2003), Consumer patronage and risk perceptions in Internet shopping, *Journal of Business Research,* 56(11), 867-875.

Ghose, Anindya and Panagiotis G. Ipeirotis (2011), Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics, *Knowledge and Data Engineering, IEEE Transactions,* 23(10), 1498-1512.

Godes, David and Dina Mayzlin (2004), Using online conversations to study word-of-mouth communication, *Marketing Science*, 23(4), 545-560

Gudykunst, William B., and Stella Ting-Toomey (1988), *Culture and interpersonal communication*. Newbury Park, CA: Sage.

Gudykunst, William B., Yuko Matsumoto, Stella Ting-Toomey, Tsukasa Nishida, Kwangsu Kim, and Sam Heyman (1996), The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures, *Human Communication Research,* 22(4), 510-543.

Hall, E. (1976), *Beyond Culture*. New York: Doubleday.

Hann, Il-Horn and Christian Terwiesch (2003), Measuring the frictional costs of online transaction: The case of a name-your-own-price channel, *Management Science*, 49(11), 1563-1579.

Hermeking, Marc (2005), Culture and internet consumption: contributions from cross-cultural marketing and advertising research, *Journal of Computer‐Mediated Communication,* 11(1), 192-216.

Herr, Paul M., Frank R. Kardes, and John Kim (1991), Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective, *Journal of Consumer Research,* 454-462.

Hofstede, G. (1980), *Culture's Consequences,* Beverly Hills, CA: Sage.

Hofstede, G. (2001), Cultures's Consequences: Comparing values, behaviors, institutions and organizations across nations, Ed. Geert Hofstese, Sage.

Hu, Nan, Jie Zhang, and Paul A. Pavlou (2009), Overcoming the J-shaped distribution of product reviews, *Communications of the ACM,* 52(10), 144-147.

Huang, Jen-Hung and Yi-Fen Chen (2006), Herding in online product choice, *Psychology & Marketing*, 23(5), 413-428.

Huang, Peng, Nicholas H. Lurie, and Sabyasachi Mitra (2009), Searching for experience on the web: an empirical examination of consumer behavior for search and experience goods, *Journal of Marketing,* 73(2), 55-69.

Hult, G. Tomas M., David J. Ketchen, David A. Griffith, Carol A. Finnegan, Tracy Gonzalez-Padron, Nukhet Harmancioglu, Ying Huang, M. Berk Talay, and S. Tamer Cavusgil (2008), Data equivalence in cross-cultural international business research: assessment and guidelines, *Journal of International Business Studies,* 39(6), 1027-1044.

Hunt, Earl, and Franca Agnoli (1991), The Whorfian hypothesis: A cognitive psychology perspective, *Psychological Review,* 98(3), 377.

Janiszewski, Chris (1993), Preattentive mere exposure effects, *Journal of Consumer Research,* 376-392.

Jansen, Bernard J. and Amanda Spink (2006), How are we searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs, *Information Processing and Management,* 42(1), 248-263.

Johanson, Jan, and Jan-Erik Vahlne (1997), The internationalization process of the firm-a model of knowledge development and increasing foreign market commitments, *Journal of International Business Studies,* 23-32.

Joshi, A. & Lahiri, N. (2012), Knowledge transfer and cross-border alliance formation. *Working Paper*.

Johnson, Timothy, Patrick Kulesa, Young Ik Cho, and Sharon Shavitt (2005), The relation between culture and response styles evidence from 19 countries, *Journal of Cross-cultural Psychology,* 36(2), 264-277.

Khare, Adwait, Lauren I. Labrecque, and Anthony K. Asare (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.

Kim, Donghoon, Yigang Pan, and Heung Soo Park (1998), High-versus low-context culture: A comparison of Chinese, Korean, and American cultures, *Psychology and Marketing,* 15(6), 507-521.

Kim, Jiyeon, and Sandra Forsythe (2008), Adoption of Virtual Try-on technology for online apparel shopping, *Journal of Interactive Marketing,* 22(2), 45-59.

Kim, Junyong, and Pranjal Gupta (2012), Emotional expressions in online user reviews: How they influence consumers' product evaluations, *Journal of Business Research,* 65(7), 985-992.

King, Robert Allen, Pradeep Racherla, and Victoria D. Bush (2014), What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature, *Journal of Interactive Marketing,* 28(3), 167-183.

Kozinets, Robert V. (1999), E-tribalized marketing? The strategic implications of virtual communities of consumption, *European Management Journal*, 17(3), 252-64.

Kozinets, Robert V. (2002), The field behind the screen: Using netnography for marketing research in online communities, *Journal of Marketing Research*, 39(1), 61-72.

Kozinets, Robert V., Kristine de Valck, Andreas C. Wojnicki, and Sarah J. S. Wilner (2010), Networked narratives: Understanding word-of-mouth marketing in online communities, *Journal of Marketing*, 74(2), 71-89.

Krosnick, Jon A., David S. Boninger, Yao C. Chuang, Matthew K. Berent, and Catherine G. Carnot (1993), Attitude strength: One construct or many related constructs?, *Journal of Personality and Social Psychology,* 65(6), 1132.

Kumar, Nanda, and Izak Benbasat (2006). Research note: the influence of recommendations and consumer reviews on evaluations of websites, *Information Systems Research,* 17(4), 425-439.

Kumar, V., Lerzan Aksoy, Bas Donkers, Rajkumar Venkatesan, Thorston Wiesel and Sebastian Tillmanns (2010), Undervalued or overvalued customers: Capturing total customer engagement value, *Journal of Service Research,* 13(3), 297-310.

Labroo, Aparna A., Ravi Dhar, and Norbert Schwarz (2008), Of frog wines and frowning watches: Semantic priming, perceptual fluency, and brand evaluation, *Journal of Consumer Research,* 34(6), 819-831.

Landsman, S. (2013), Love it or leave it – Growing power of customer reviews. June 9. http://www.cnbc.com/id/100792646

Lau-Gesk, Loraine, and Joan Meyers-Levy (2009), Emotional persuasion: When the valence versus the resource demands of emotions influence consumers attitudes, *Journal of Consumer Research*, 36(4), 585-599.

Lee, Angela Y. (2002), Effects of implicit memory on memory-based versus stimulus-based brand choice, *Journal of Marketing Research,* 39(4), 440-454.

Lee, Angela Y., and Aparna A. Labroo (2004), The effect of conceptual and perceptual fluency on brand evaluation, *Journal of Marketing Research,* 41(2), 151-165.

Liao, H., 2013. How much can one express in 140 characters? Comparison between English and other languages like Chinese. *2013 Chinese Internet Research Conference.* http://people.oii.ox.ac.uk/hanteng/2013/04/16/how-much-can-one-express-in-140-characters-comparison-between-english-and-other-languages-like-chinese/comment-page-1/

Lim, Kai H., Kwok Leung, Choon Ling Sia, and Matthew KO Lee (2004), Is eCommerce boundary-less? Effects of individualism–collectivism and uncertainty avoidance on Internet shopping, *Journal of International Business Studies,* 35(6), 545-559.

Liu, Jingjing, Yunbo Cao, Chin-Yew Lin, Yalou Huang, and Ming Zhou (2007), Low-quality product review detection in opinion summarization*, Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague, Czech Republic, 334-342.

Liu, Raymond R., and Peter McClure (2001), Recognizing cross-cultural differences in consumer complaint behavior and intentions: an empirical examination, *Journal of Consumer Marketing,* 18(1), 54-75.

Liu, Yong (2006), Word of mouth for movies: Its dynamics and impact on box office revenue, *Journal of Marketing*, 70(3), 74-89.

Lohmann, Johannes (2011), Do language barriers affect trade?, *Economics Letters,* 110(2), 159-162.

Luca, Michael (2011), Reviews, reputation, and revenue: The case of Yelp. com.*Com (September 16, 2011), Harvard Business School NOM Unit Working Paper*, 12-016.

Ludwig, Stephan, Ko de Ruyter, Mike Friedman, Elisabeth C. Bruggen, Martin Wetzels, and Gerard Pfann (2013), More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates, *Journal of Marketing*, 77(1), 87-103.

Luo, Yadong, and Oded Shenkar (2006), The multinational corporation as a multilingual community: Language and organization in a global context, *Journal of International Business Studies,* 37(3), 321-339.

MacInnis, Deborah J (2011), A framework for conceptual contributions in marketing, *Journal of Marketing,* 75(4), 136-154.

Mathwick, Charla (2002), Understanding the online consumer: A typology of online relational norms and behavior, *Journal of Interactive Marketing*, 16(1), 40-55.

Menon, Geeta, and Priya Raghubir (2003), Ease-of-Retrieval as an Automatic Input in Judgments: A Mere-Accessibility Framework?, *Journal of Consumer Research,* 30(2), 230-243.

Miller, George A. (1956), The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychological Review,* 63(2), 81

Mita, Theodore, Marshal Dermer, and Jeffrey Knight (1977), Reversed facial images and the mere-exposure hypothesis, *Journal of Personality and Social Psychology*, 35(8), 597-601.

Moe, Wendy W. and David A. Schweidel (2011), Online product opinions: Incidence, Evaluation, and Evolution, *Marketing Science,* 31(3), 372-386.

Moe, Wendy W. and Michael Trusov (2011), The value of social dynamics in online product ratings forums, *Journal of Marketing Research, 48*(3), 444-456.

Money, R. Bruce, Mary C. Gilly, and John L. Graham (1998), Explorations of national culture and word-of-mouth referral behavior in the purchase of industrial services in the United States and Japan, *The Journal of Marketing,* 76-87.

Moon, Jae Yun and Lee S. Sproull (2008), The role of feedback in managing the internet-based volunteer workforce, *Information Systems Research*, 19(4), 494-515.

Mudambi, Susan M. and David Schuff (2010), What makes a helpful online review? A study of customer reviews on amazon.com, *MIS Quarterly,* 34(1), 185-200.

Naylor, Rebecca Walker, Cait Poynor Lamberton, and David A. Norton (2011), Seeing ourselves in others: Reviewer ambiguity, egocentric anchoring, and persuasion, *Journal of Marketing Research, 48*(3), 617-631.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko (2012), Mine your own business: Market-structure surveillance through text mining, *Marketing Science,* 31(3), 521-543.

Neubig, Graham, and Kevin Duh (2013), How Much Is Said in a Tweet? A Multilingual, Information-theoretic Perspective, *AAAI Spring Symposium: Analyzing Microtext*.

Ngai, Eric WT, Vincent CS Heung, Y. H. Wong, and Fanny KY Chan (2007), Consumer complaint behaviour of Asians and non-Asians about hotel services: an empirical analysis, *European Journal of Marketing,* 41(11/12), 1375-1391.

Ngo-Ye, Thomas L., and Atish P. Sinha (2014), The influence of reviewer engagement characteristics on online review helpfulness: A text regression model, *Decision Support Systems,* 61, 47-58.

Nielsen, 2012. Global consumers trust in earned advertising grows in importance. http://www.nielsen.com/content/corporate/us/en/press-room/2012/nielsen-global-consumers-trust-in-earned-advertising-grows.html Accessed 10 October 2014.

Nielsen, 2014. Mobile Millennials: Over 85% of generation Y owns smartphones. http://www.nielsen.com/us/en/insights/news/2014/mobile-millennials-over-85-percent-of-generation-y-owns-smartphones.html Accessed 5 May 2015.

NIST (2006), Machine Translation Evaluation Official Results. http://www.itl.nist.gov/iad/mig/tests/mt/2006/doc/mt06eval_official_results.html. Accessed 23 July 2013.

Nisbett, Richard. E., Kaiping Peng, Incheol Choi, I., and Ara Norenzayan (2001), Culture and systems of thought: Holistic vs. analytic cognition, *Psychological Review*, 108(2), 291-310.

Novemsky, Nathan, Ravi Dhar, Norbert Schwarz, and Itamar Simonson (2007), Preference fluency in choice, *Journal of Marketing Research,* 44(3), 347-356.

Pan, Bing, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gary, and Laura Granka, (2007), "In Google we trust: Users' decisions on rank, position, and relevance," *Journal of Computer-Mediated Communication,* 12(3), 801-823.

Pan, Bing, Helene A. Hembrooke, Geri K. Gay, Laura Granka, Matthew K. Feusner, and Jill K. Newman (2004), The determinants of web page viewing behavior: An eye-tracking study, P*roceedings of the Eye Tracking Research & Applications Symposium on Eye Tracking Research & Applications,* San Antonio, Texas, 147-154.

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

Park, Cheol, and Thae Min Lee (2009), Antecedents of online reviews' usage and purchase influence: an empirical comparison of US and Korean consumers, *Journal of Interactive Marketing,* 23(4), 332-340.

Pennebaker, James W., Cindy K. Chung, Molly Ireland, Amy Gonzales, and Roger J. Booth, (2007), The development and psychometric properties of LIWC2007, *LIWC.Net, 2007.*

Prentice, Deborah A., Dale T. Miller, and Jenifer R. Lightdale (1994), Asymmetries in attachments to groups and to their members: Distinguishing between common-identity and common-bond groups, *Personality and Social Psychology Bulletin*, 29(5), 484-493.

Qiu, Lingyun, Jun Pang, and Kai H. Lim (2012), Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence, *Decision Support Systems,* 54(1), 631-643.

Reber, Rolf, Norbert Schwarz, and Piotr Winkielman (2004), Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?, *Personality and Social Psychology Review*, 8(4), 364-382.

Reber, Rolf, Poitr Winkielman, and Norbert Schwarz (1998), Effects of perceptual fluency on affective judgments, *Psychological Science*, 9(1), 45-48.

Riefler, Petra, Adamantios Diamantopoulos, and Judy A. Siguaw (2012), Cosmopolitan consumers as a target group for segmentation, *Journal of International Business Studies,4*3(3), 285-305.

Reis, Harry T., Michael R. Maniaci, Peter A. Caprariello, Paul W. Eastwick, and Eli J. Finkel (2011), Familiarity does indeed promote attraction in live interaction, *Journal of Personality and Social Psychology,* 101(3), 557.

Ren, Yuqing, Robert Kraut and Sara Kiesler (2007), Applying common identity and bond theory to design of online communities, *Organization Studies*, 28, 377-408.

Rill, Sven, Nikolaos Korfiatis, Jorg Scheidt, and Roberto Zicari (2013), Mining Negativity from online reviews: A comparison between search and experience goods, *Available at SSRN: http://ssrn.com/abstract=2199626 or http://dx.doi.org/10.2139/ssrn.2199626*

Rosen, R. (2012), How much can you say in 140 characters? A lot, if you speak Japanese. *The Atlantic*, September 16.

Roster, Catherine, Gerald Albaum, and Robert Rogers (2006), Can cross-national/cultural studies presume etic equivalency in responclents' use of extreme categories of Likert rating scales?, *International Journal of Market Research,* 48(6), 741-759.

Salfino, Catherine (2014), Retail goes virtual to perfect the fit, (accessed July 14, 2014) [available at https://www.sourcingjournalonline.com/retail-goes-virtual-perfect-fit-salfinotd/]

Schau, Hope, Albert Muniz Jr., and Eric Arnould (2009), How brand community practices create value, *Journal of Marketing*, 73(5), 30-51.

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

Schlosser, Ann E. (2005), Posting versus lurking: Communicating in a multiple audience context, *Journal of Consumer Research*, 32(2), 260-265.

Schlosser, Ann 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.

Schwartz, Shalom H. (1999), A theory of cultural values and some implications for work, *Applied Psychology,* 48(1), 23-47.

Sen, Shahana and Dawn 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.

Seraj, Mina (2012), We create, we connect, we respect, therefore we are: intellectual, social, and cultural value in online communities, *Journal of Interactive Marketing,* 26(4), 209-222.

Sia, Choon Ling, Kai H. Lim, Kwok Leung, Matthew KO Lee, Wayne Wei Huang, and Izak Benbasat (2009), Web strategies to promote internet shopping: is cultural-customization needed?, *MIS Quarterly,* 491-512.

Singh, Jagdip (1988), Consumer complaint intentions and behavior: definitional and taxonomical issues, *The Journal of Marketing,* 93-107.

Singh, Nitish, Vikas Kumar, and Daniel Baack (2005), Adaptation of cultural content: evidence from B2C e-commerce firms, *European Journal of Marketing,* 39(1/2), 71-86.

Smith, Anne M., and Nina L. Reynolds (2002), Measuring cross-cultural service quality: a framework for assessment, *International Marketing Review,* 19(5), 450-481.

Spencer-Rodgers, Julie, Kaiping Peng, Lei Wang, and Yubo Hou (2004), Dialectical self-esteem and East-West differences in psychological well-being, *Personality and Social Psychology Bulletin,* 30(11), 1416-1432.

Spencer-Rodgers, Julie, Melissa J. Williams, and Kaiping Peng (2010), Cultural differences in expectations of change and tolerance for contradiction: A decade of empirical research, *Personality and Social Psychology Review*.

Spink, Amanda, Dietmar Wolfram, Major B. J. Jansen and Tefko Saracevic (2001), Searching the Web: The public and their queries, *Journal of the American Society for Information Science and Technology,* 52(3), 226-234.

Sridhar, Shrihari and Raji Srinivasan (2012), Social influence effects in online product ratings, *Journal of Marketing*, 76(3), 70-88.

Steenkamp, Jan-Benedict EM, Rajeev Batra, and Dana L. Alden (2003), How perceived brand globalness creates brand value, *Journal of International Business Studies,* 34(1), 53-65.

Steenkamp, Jan-Benedict EM, and Hans Baumgartner (1998), Assessing measurement invariance in cross-national consumer research, *Journal of Consumer Research,* 25(1), 78-107.

Steenkamp, Jan-Benedict EM, and Frenkel Ter Hofstede (2002), International market segmentation: issues and perspectives, *International Journal of Research in Marketing,* 19(3), 185-213.

Stewart, David W. and Paul A. Pavlou (2002), From consumer response to active consumer: Measuring the effectiveness of interactive media, *Journal of the Academy of Marketing Science*, 30(4), 376-396).

Stout, Hilary (2014), Mirror, mirror in the app: What's the fairest shade and shadow of them all? (accessed June 25, 2014) [available at http://www.nytimes.com/2014/05/15/business/mirror-mirror-in-the-app-whats-the-fairest-shade-of-all.html?hpw&rref=business&_r=1]

Summers, B. (2010), What's the equivalent of Twitter's 140 character limit for non-Latin character sets? *Ben Summers' blog*. Retrieved May 1, 2014, from http://bens.me.uk/2010/twitter-charset-experiment

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

Sundar, S. Shyam and Clifford Nass (2001), Conceptualizing sources in online news, *Journal of Communication,* 51(1), 52-72.

Tang, Tanya, Eric Fang, and Feng Wang (2014), Is neutral really neutral? The effects of neutral user-generated content on product sales, *Journal of Marketing* 78(4), 41-58.

Tannen, Deborah (1984), The pragmatics of cross-cultural communication, *Applied Linguistics,* 5(3), 189-195.

Tellis, Gerard J., and Deepa Chandrasekaran (2010), Extent and impact of response biases in cross-national survey research, *International Journal of Research in Marketing,* 27(4), 329-341.

Thomas, Manoj, and Vicki G. Morwitz (2009), The ease-of-computation effect: The interplay of metacognitive experiences and naive theories in judgments of price differences, *Journal of Marketing Research,* 46(1), 81-91.

Tian, Robert G., and Charles Emery (2002), Cross-cultural issues in Internet marketing, *Journal of American Academy of Business,* 1(2), 217-224.

Tirunillai, Seshadri, and Gerard J. Tellis (2012), Does chatter really matter? Dynamics of user-generated content and stock performance, *Marketing Science,* 31(2), 198-215.

Triandis, Harry C. (1988), Collectivism vs. individualism: A reconceptualization of a basic concept in cross-cultural psychology, *Cross-cultural Studies of Personality, Attitudes, and Cognition,* 60-95.

Usunier, Jean-Claude, and Nicolas Roulin (2012), The influence of high-and low-context communication styles on the design, content, and language of business-to-business web sites, *Journal of Business Communication,* 47(2), 189-227.

Van Herk, Hester, Ype H. Poortinga, and Theo MM Verhallen (2004), Response styles in rating scales evidence of method bias in data from six EU countries, *Journal of Cross-Cultural Psychology,* 35(3), 346-360.

Varadarajan, Rajan, Raji Srinivasan, Gautham Gopal Vadakkepatt, Manjit S. Yadav, Paul A. Pavlou, Sandeep Krishnamurthy, and Tom Krause (2010), Interactive technologies and retailing strategy: a review, conceptual framework and future research directions, *Journal of Interactive Marketing,* 24(2), 96-110.

Vargo, Stephen L., and Robert F. Lusch (2004), Evolving to a new dominant logic for marketing, *Journal of Marketing,* 68(1), 1-17.

Vishwanath, Arun (2003), Comparing online information effects a cross-cultural comparison of online information and uncertainty avoidance, *Communication Research,* 30(6), 579-598.

W3Techs 2014. Usage of content languages for websites. http://w3techs.com/technologies/overview/content_language/all Accessed 11 December 2014

Wang, Qiuzhen, Yi Yang, Qi Wang, and Qingguo Ma (2014), The effect of human image in B2C website design: an eye-tracking study, *Enterprise Information Systems,* 8(5), 582-605.

Whorf, Benjamin Lee, Penny Lee, Stephen C. Levinson, and John B. Carroll (2012), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf,* MIT Press.

Wilcox, Keith, and Sangyoung Song (2011), Discrepant fluency in self-customization, *Journal of Marketing Research,* 48(4), 729-740.

Wu, Philip Fei (2013), In search of negativity bias: An empirical study of perceived helpfulness of online reviews, *Psychology & Marketing,* 30(11), 971-984.

Würtz, Elizabeth (2005), Intercultural communication on Web sites: a cross-cultural analysis of Web sites from high-context cultures and low-context cultures, *Journal of Computer‐Mediated Communication,* 11(1), 274-299.

Yadav, Manjit S., and Paul A. Pavlou (2014), Marketing in Computer-Mediated Environments: Research Synthesis and New Directions, *Journal of Marketing,* 78(1), 20-40.

Yadav, Manjit S. and Rajan Varadarajan (2005), Interactivity in the electronic marketplace: An exposition of the concept and implications for research, *Journal of the Academy of Marketing Science*, 33(4), 585-603.

Yang, Xiaojing, Jing Zhang, and Laura A. Peracchio (2010), Understanding the impact of self-concept on the stylistic properties of images, *Journal of Consumer Psychology,* 20(4), 508-520.

Yin, Dezhi, Samuel Bond, and Han Zhang (2014), Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews, *MIS Quarterly* 38(2), 539-560.

Zajonc, Robert B. (1968), Attitudinal effects of mere exposure, *Journal of Personality and Social Psychology,* 9(2), 1.

Zhang, Jason Q., Georgiana Craciun and Dongwoo Shin (2010), When does electronic word-of-mouth matter? A study of consumer product reviews, *Journal of Business Research*, 63, 1336-1341.

Zhao, Yi, Sha Yang, Vishal Narayan, and Ying Zhao (2013), Modeling consumer learning from online product reviews, *Marketing science,* 32(1), 153-169.

Zhu, Feng and Xiaoquan (Michael) Zhang (2010), Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics, *Journal of Marketing, 74*(2), 133-148.

APPENDIX A: INSTRUCTIONS FOR NORM VIOLATIONS

Thank you for your assistance with this study. The purpose of your involvement is to evaluate a set of book reviews from Amazon.com according to certain guidelines. Amazon.com sets forth review guidelines and rules for the posting of online reviews and systematically removes gross violations such as obscenity and commercial promotion.

More subtle violations of Amazon.com review content guidelines are typically not removed from the site, but still violate what is generally considered an appropriate review. Chief among these violations are reviews that contain "off-topic" information. Put simply, reviews where the primary focus is not evaluation of the book being reviewed. Please review the guidelines below for examples of situations where a review would be considered inappropriate:

- Review content primarily evaluates the seller rather than the book purchased. Prime examples include but are not limited to:
    - Customer service received from the seller
    - Shipping concerns
    - Electronic delivery issues
    - Pricing concerns
    - Web-interface issues
    - Purchase page or product listing

- Comments that primarily address or respond to the review of another reviewer

- Content that primarily expresses the reviewer's opinions of subjects unrelated to the content of the book being reviewed. Prime examples include but are not limited to:
    - Unrelated political and religious views
    - Unrelated discussions of personal habits and tendencies
    - Grossly insincere comments, or content unrelated to the book being reviewed

Your task is to read a set of reviews and determine in your opinion whether or not each review is appropriate based on the guidelines provided above. You have been provided with a text document containing the reviews for evaluation and a spreadsheet to record your answers. If you think the review violates the guidelines described here, mark the response column with a "Y". If you think the review is appropriate and does not violate any of the guidelines presented above, mark the response column with an "N".

APPENDIX B: PRODUCT LIST FOR DEVELOPMENT OF THE CHARACTER COUNT LANGUAGE INDEX (CCLI)

Country	Category	Product
China	Electronics	Apple iPod Touch 4 32GB
	Books	*The Art and Science of Personal Magnetism* by Theron Q. Dumont
	Cameras	Canon Powershot G12
France	Electronics	Apple iPod Touch 4 32GB
	Books	*The Girl With the Dragon Tattoo* by Stieg Larsson
	Cameras	Canon PowerShot SX130
		Canon PowerShot SX210
		Canon PowerShot SX30
Germany	Electronics	Apple iPod Touch 4 32GB
	Books	*The Girl With the Dragon Tattoo* by Stieg Larsson
	Cameras	Canon PowerShot SX210
Japan	Electronics	Apple iPod Touch 4 32GB
	Books	*The Girl Who Kicked the Hornet's Nest* by Stieg Larsson
	Cameras	Canon PowerShot SX130IS
United States	Electronics	Apple iPod Touch 4 8GB
	Books	*The Girl Who Kicked the Hornet's Nest* by Stieg Larsson
	Cameras	Canon PowerShot SD 1300IS
Spain	Electronics	Apple iPod Nano 7 16 GB
	Books	*Espia de Dios* by Juan Gomez-Jurado
	Cameras	Canon PowerShot SX50 HS

APPENDIX C: VALIDITY CHECK OF GOOGLE TRANSLATE

To verify that Google Translate produced translations with character counts equivalent to native speakers, a sample of 60 reviews used to create the CCLI was gathered. For each language (French, German, Japanese, Chinese, and Spanish), four camera reviews, four iPod reviews, and four book reviews were gathered. Native-speaking and bilingual coders manually translated each native language review into English. The average character count of the translations provided by the native speakers was compared to the character counts produced by Google Translate using paired t-tests. As shown in the table below there were no significant difference between the manual human translations and the Google Translate-generated text. From this analysis it can be concluded that, for the purposes of this study, Google Translate provides the basis for a reasonable estimate of character counts.

Comparison of Machine vs. Human Translation Character Counts

Language (From → To)	Mean Character Count – Google	Mean Character Counts - Human	t-value	Sig.
Languages in Amazon Sample				
French to English	824.08	831.29	-0.71	0.495
German to English	1044.42	1036.00	0.69	0.502
Japanese to English	1205.17	1191.17	0.42	0.685
Chinese to English	842.67	926.00	-1.57	0.146
Spanish to English	1164.50	1180.45	-0.56	0.583

N=60 total
(12 per analysis)