MIS Quarterly

WHAT MAKES A HELPFUL ONLINE REVIEW? A STUDY OF CUSTOMER REVIEWS ON AMAZON.COM[1]

By: **Susan M. Mudambi**
Department of Marketing and Supply
  Chain Management
Fox School of Business
Temple University
524 Alter Hall
1801 Liacouras Walk
Philadelphia, PA 19122
U.S.A.
Susan.Mudambi@temple.edu

David Schuff
Department of Management Information Systems
Fox School of Business
Temple University
207G Speakman Hall
1810 North 13th Street
Philadelphia, PA 19122
U.S.A.
David.Schuff@temple.edu

Abstract

Customer reviews are increasingly available online for a wide range of products and services. They supplement other information provided by electronic storefronts such as product descriptions, reviews from experts, and personalized advice generated by automated recommendation systems. While researchers have demonstrated the benefits of the presence of customer reviews to an online retailer, a largely uninvestigated issue is what makes customer reviews helpful

to a consumer in the process of making a purchase decision. Drawing on the paradigm of search and experience goods from information economics, we develop and test a model of customer review helpfulness. An analysis of 1,587 reviews from Amazon.com across six products indicated that review extremity, review depth, and product type affect the perceived helpfulness of the review. Product type moderates the effect of review extremity on the helpfulness of the review. For experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings. For both product types, review depth has a positive effect on the helpfulness of the review, but the product type moderates the effect of review depth on the helpfulness of the review. Review depth has a greater positive effect on the helpfulness of the review for search goods than for experience goods. We discuss the implications of our findings for both theory and practice.

Introduction

As consumers search online for product information and to evaluate product alternatives, they often have access to dozens or hundreds of product reviews from other consumers. These customer reviews are provided in addition to product descriptions, reviews from experts, and personalized advice generated by automated recommendation systems. Each of these options has the potential to add value for a prospective customer. Past research has extensively examined the role of expert reviews (Chen and Xie 2005), and the role of online recommendation systems (Bakos 1997; Chen et al. 2004; Gretzel and Fesenmaier 2006), and the positive effect feedback mechanisms can have on buyer trust (Ba and Pavlou

---

2002; Pavlou and Gefen 2004). More recently, research has examined the role of online customer product reviews, specifically looking at the characteristics of the reviewers (Forman et al. 2008, Smith et al. 2005) and self-selection bias (Hu et al. 2008; Li and Hitt 2008). Recent research has also shown that customer reviews can have a positive influence on sales (see Chen et al. 2008; Chevalier and Mayzlin 2006; Clemons et al. 2006; Ghose and Ipeirotis 2006). Specifically, Clemons et al. (2006) found that strongly positive ratings can positively influence the growth of product sales, and Chen et al. (2008) found that the quality of the review as measured by helpfulness votes also positively influences sales. One area in need of further examination is what makes an online review helpful to consumers.

Online customer reviews can be defined as *peer-generated product evaluations posted on company or third party websites*. Retail websites offer consumers the opportunity to post product reviews with content in the form of numerical star ratings (usually ranging from 1 to 5 stars) and open-ended customer-authored comments about the product. Leading online retailers such as Amazon.com have enabled consumers to submit product reviews for many years, with other retailers offering this option to consumers more recently. Some other firms choose to buy customer reviews from Amazon.com or other sites and post the reviews on their own electronic storefronts. In this way, the reviews themselves provide an additional revenue stream for Amazon and other online retailers. A number of sites that provide consumer ratings have emerged in specialty areas (Dabholkar 2006) such as travel (www.travelpost.com) and charities (www.charitynavigator.org).

The presence of customer reviews on a website has been shown to improve customer perception of the usefulness and social presence of the website (Kumar and Benbasat 2006). Reviews have the potential to attract consumer visits, increase the time spent on the site ("stickiness"), and create a sense of community among frequent shoppers. However, as the availability of customer reviews becomes widespread, the strategic focus shifts from the mere *presence* of customer reviews to the customer evaluation and use of the reviews. Online retailers have an incentive to provide online content that customers perceive to be valuable, and sites such as eOpinions and Amazon.com post detailed guidelines for writing reviews. Making a better decision more easily is the main reason consumers use a ratings website (Dabholkar 2006), and the perceived diagnosticity of website information positively affects consumers' attitudes toward shopping online (Jiang and Benbasat 2007).

Online retailers have commonly used review "helpfulness" as the primary way of measuring how consumers evaluate a re-

view. For example, after each customer review, Amazon.com asks, "Was this review helpful to you?" Amazon provides helpfulness information alongside the review ("26 of 31 people found the following review helpful") and positions the most helpful reviews more prominently on the product's information page. Consumers can also sort reviews by their level of helpfulness. However, past research has not provided a theoretically grounded explanation of what constitutes a helpful review. We define a helpful customer review as *a peer-generated product evaluation that facilitates the consumer's purchase decision process*.

Review helpfulness can be seen as a reflection of review diagnosticity. Interpreting helpfulness as a measure of perceived value in the decision-making process is consistent with the notion of information diagnosticity found in the literature (see Jiang and Benbasat 2004 2007; Kempf and Smith, 1998; Pavlou and Fygenson 2006; Pavlou et al. 2007). Customer reviews can provide diagnostic value across multiple stages of the purchase decision process. The purchase decision process includes the stages of need recognition, information search, evaluation of alternatives, purchase decision, purchase, and post-purchase evaluation (adapted from Kotler and Keller 2005). Once a need is recognized, consumers can use customer reviews for information search and the evaluation of alternatives. The ability to explore information about alternatives helps consumers make better decisions and experience greater satisfaction with the online channel (Kohli et al. 2004). For some consumers, information seeking is itself a source of pleasure (Mathwick and Rigdon 2004). After the purchase decision and the purchase itself, some consumers return to the website in the post-purchase evaluation stage to post comments on the product purchased. After reading peer comments, consumers may become aware of an unfilled product need, thereby bringing the purchase decision process full circle.

This implies that online retail sites with more helpful reviews offer greater potential value to customers. Providing easy access to helpful reviews can create a source of differentiation. In practice, encouraging quality customer reviews does appear to be an important component of the strategy of many online retailers. Given the strategic potential of customer reviews, we draw on information economics theory and on past research to develop a conceptual understanding of the components of helpfulness. We then empirically test the model using actual customer review data from Amazon.com. Overall, the analysis contributes to a better understanding of what makes a customer review helpful in the purchase decision process. In the final section, we conclude with a discussion of the managerial implications.

Theoretical Foundation and Model ▰

The economics of information provides a relevant foundation to address the role of online customer reviews in the consumer decision process. Consumers often must make purchase decisions with incomplete information as they lack full information on product quality, seller quality, and the available alternatives. They also know that seeking this information is costly and time consuming, and that there are trade-offs between the perceived costs and benefits of additional search (Stigler 1961). Consumers follow a purchase decision process that seeks to reduce uncertainty, while acknowledging that purchase uncertainty cannot be totally eliminated.

Therefore, the total cost of a product must include both the product cost and the cost of search (Nelson 1970). Both physical search and cognitive processing efforts can be considered search costs. For a wide range of choices, consumers recognize that there are tradeoffs between effort and accuracy (Johnson and Payne 1985). Those who are willing to put more effort into the decision process expect, at least partially, increased decision accuracy. Consumers can use decision and comparison aids (Todd and Benbasat 1992) and numerical content ratings (Poston and Speier 2005) to conserve cognitive resources and reduce energy expenditure, but also to ease or improve the purchase decision process. One such numerical rating, the star rating, has been shown to serve as a cue for the review content (Poston and Speier 2005).

A key determinant of search cost is the nature of the product under consideration. According to Nelson (1970, 1974), *search goods* are those for which consumers have the ability to obtain information on product quality prior to purchase, while *experience goods* are products that require sampling or purchase in order to evaluate product quality. Examples of search goods include cameras (Nelson 1970) and natural supplement pills (Weathers et al. 2007), and examples of experience goods include music (Bhattacharjee et al. 2006; Nelson 1970) and wine (Klein 1998). Although many products involve a mix of search and experience attributes, the categorization of search and experience goods continues to be relevant and widely accepted (Huang et al. 2009). Products can be described as existing along a continuum from pure search goods to pure experience goods.

To further clarify the relevant distinctions between search and experience goods, the starting point is Nelson's (1974, p. 738) assertion that "goods can be classified by whether the quality variation was ascertained predominantly by search or by experience." Perceived quality of a search good involves attributes of an objective nature, while perceived quality of an experience good depends more on subjective attributes that are a matter of personal taste. Several researchers have focused on the differing information needs of various products and on how consumers evaluate and compare their most relevant attributes. The dominant attributes of a search good can be evaluated and compared easily, and in an objective manner, without sampling or buying the product, while the dominant attributes of an experience goods are evaluated or compared more subjectively and with more difficulty (Huang et al. 2009). Unlike search goods, experience goods are more likely to require sampling in order to arrive at a purchase decision, and sampling often requires an actual purchase. For example, the ability to listen online to several 30-second clips from a music CD allows the customer to gather pre-purchase information and even attain a degree of "virtual experience" (Klein 1998), but assessment of the full product or the full experience requires a purchase. In addition, Weathers et al. (2007) categorized goods according to whether or not it was necessary to go beyond simply reading information to also use one's senses to evaluate quality.

We identify an *experience good* as one in which it is relatively difficult and costly to obtain information on product quality prior to interaction with the product; key attributes are subjective or difficult to compare, and there is a need to use one's senses to evaluate quality. For a *search good*, it is relatively easy to obtain information on product quality prior to interaction with the product; key attributes are objective and easy to compare, and there is no strong need to use one's senses to evaluate quality.

This difference between search and experience goods can inform our understanding of the helpfulness of an online customer review. Customer reviews are posted on a wide range of products and services, and have become part of the decision process for many consumers. Although consumers use online reviews to help them make decisions regarding both types of products, it follows that a purchase decision for a search good may have different information requirements than a purchase decision for an experience good.

In the economics of information literature, a close connection is made between information and uncertainty (Nelson 1970). Information quality is critical in online customer reviews, as it can reduce purchase uncertainty. Our model of customer review helpfulness, as illustrated in Figure 1, starts with the assumption of a consumer's need to reduce purchase uncertainty. Although previous research has analyzed both product and seller quality uncertainty (Pavlou et al. 2007), we examine the helpfulness of reviews that focus on the product itself, not on reviews of the purchase experience or the seller.

In past research on online consumers, diagnosticity has been defined and measured in multiple ways, with a commonality of the helpfulness to a decision process, as subjectively perceived by consumers. Kempf and Smith (1998) assessed overall product-level diagnosticity by asking how helpful the website experience was in judging the quality and performance of the product. Product diagnosticity is a reflection of how helpful a website is to online buyers for evaluating product quality (Pavlou and Fygenson 2006; Pavlou et al. 2007). Perceived diagnosticity has been described as the perceived ability of a Web interface to convey to customers relevant product information that helps them in understanding and evaluating the quality and performance of products sold online (Jiang and Benbasat 2004, and has been measured as whether it is "helpful for me to evaluate the product," "helpful in familiarizing me with the product," and "helpful for me to understand the product" (Jiang and Benbasat 2007, p. 468).

This established connection between perceived diagnosticity and perceived helpfulness is highly relevant to the context of online reviews. For example, Amazon asks, "Was this review helpful to you?" In this context, the question is essentially an assessment of helpfulness during the product decision-making process. A review is helpful if it aids one or more stages of this process. This understanding of review helpfulness is consistent with the previously cited conceptualizations of perceived diagnosticity.

For our study of online reviews, we adapt the established view of perceived diagnosticity as perceived helpfulness to a decision process. We seek to better understand what makes a helpful review. Our model (Figure 1) illustrates two factors that consumers take into account when determining the helpfulness of a review. These are review extremity (whether the review is positive, negative, or neutral), and review depth (the extensiveness of the reviewer comments). Given the differences in the nature of information search across search and experience goods, we expect the product type to moderate the perceived helpfulness of an online customer review. These factors and relationships will be explained in more detail in the following sections.

Review Extremity and Star Ratings

Previous research on extreme and two-sided arguments raises theoretical questions on the relative diagnosticity or helpfulness of extreme versus moderate reviews. Numerical star ratings for online customer reviews typically range from one to five stars. A very low rating (one star) indicates an extremely negative view of the product, a very high rating (five stars) reflects an extremely positive view of the product,

and a three-star rating reflects a moderate view. The star ratings are a reflection of attitude extremity, that is, the deviation from the midpoint of an attitude scale (Krosnick et al. 1993). Past research has identified two explanations for a midpoint rating such as three stars out of five (Kaplan 1972; Presser and Schuman 1980). A three-star review could reflect a truly moderate review (indifference), or a series of positive and negative comments that cancel each other out (ambivalence). In either case, a midpoint rating has been shown to be a legitimate measure of a middle-ground attitude.

One issue with review extremity is how the helpfulness of a review with an extreme rating of one or five compares to that of a review with a moderate rating of three. Previous research on two-sided arguments provides theoretical insights on the relative diagnosticity of moderate versus extreme reviews. There is solid evidence that two-sided messages in advertising can enhance source credibility in consumer communications (Eisend 2006; Hunt and Smith 1987), and can enhance brand attitude (Eisend 2006). This would imply that moderate reviews are more helpful than extreme reviews.

Yet, past research on reviews provides findings with conflicting implications for review diagnosticity and helpfulness. For reviews of movies with moderate star ratings, Schlosser (2005) found that two-sided arguments were more credible and led to more positive attitudes about the movie, but in the case of movies with extreme ratings, two-sided arguments were less credible.

Other research on online reviews provides insights on the relationship between review diagnosticity and review extremity. Pavlou and Dimoka (2006) found that the extreme ratings of eBay sellers were more influential than moderate ratings, and Forman et al. (2008) found that for books, moderate reviews were less helpful than extreme reviews. One possible explanatory factor is the consumer's initial attitude. For example, Crowley and Hoyer (1994) found that two-sided arguments are more persuasive than one-sided positive arguments when the initial attitude of the consumer is neutral or negative, but not in other situations.

These mixed findings do not lead to a definitive expectation of whether extreme reviews or moderate reviews are more helpful. This ambiguity may be partly explained by the observation that previous research on moderate versus extreme reviews failed to take product type into consideration. The relative value of moderate versus extreme reviews may differ depending on whether the product is a search good or an experience good. Research in advertising has found that consumers are more skeptical of experience than search attribute claims, and more skeptical of subjective than objec-
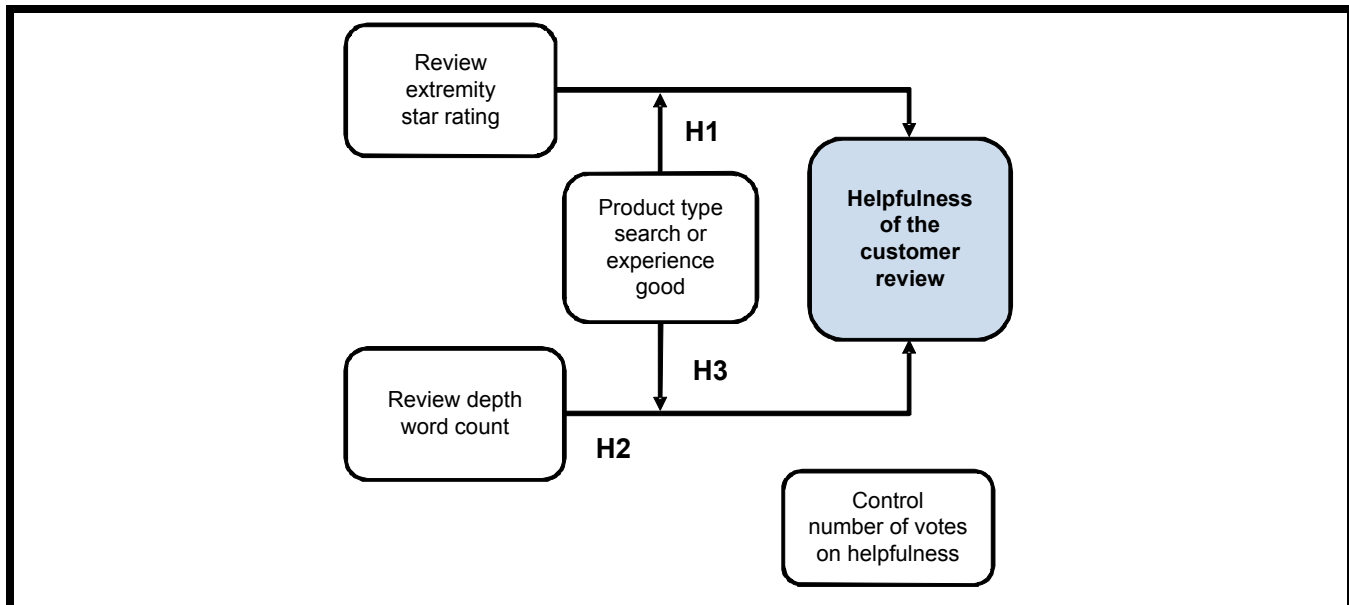
Figure 1. Model of Customer Review Helpfulness

tive claims (Ford et al. 1990). This indicates resistance to strong or extreme statements when those claims cannot be easily substantiated.

There may be an interaction between product type and review extremity, as different products have differing information needs. On consumer ratings sites, experience goods often have many extreme ratings and few moderate ratings, which can be explained by the subjective nature of the dominant attributes of experience goods. Taste plays a large role in many experience goods, and consumers are often highly confident about their own tastes and subjective evaluations, and skeptical about the extreme views of others. Experience goods such as movies and music seem to attract reviews from consumers who either love them or hate them, with extremely positive reviews especially common (Ghose and Ipeirotis 2006). Consumers may discount extreme ratings if they seem to reflect a simple difference in taste. Evidence of high levels of cognitive processing typically does not accompany extreme attitudes on experience goods. Consumers are more open to moderate ratings of experience goods, as they could represent a more objective assessment.

For experience goods, this would imply that objective content is favored, and that moderate reviews would be likely to be more helpful than either extremely negative or extremely positive reviews in making a purchase decision. For example, a consumer who has an initial positive perception of an experience good (such as a music CD) may agree with an extremely positive review, but is unlikely to find that an extreme review will help the purchase decision process. Similarly, an extremely negative review will conflict with the consumer's initial perception without adding value to the purchase decision process.

Reviews of search goods are more likely to address specific, tangible aspects of the product, and how the product performed in different situations. Consumers are in search of specific information regarding the functional attributes of the product. Since objective claims about tangible attributes are more easily substantiated, extreme claims for search goods can be perceived as credible, as shown in the advertising literature (Ford et al. 1990). Extreme claims for search goods can provide more information than extreme claims for experience goods, and can show evidence of logical argument. We expect differences in the diagnosticity and helpfulness of extreme reviews across search and experience goods. Therefore, we hypothesize

> H1. *Product type moderates the effect of review extremity on the helpfulness of the review. For experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings.*

Sample reviews from Amazon.com can serve to illustrate the key differences in the nature of reviews of experience and search goods. As presented in Appendix A, reviews with extreme ratings of experience goods often appear very subjec-

tive, sometimes go off on tangents, and can include sentiments that are unique or personal to the reviewer. Reviews with moderate ratings of experience goods have a more objective tone, keep more focused, and reflect less idiosyncratic tastes. In contrast, both extreme and moderate reviews of search goods often take an objective tone, refer to facts and measurable features, and discuss aspects of general concern. Overall, we expect extreme reviews of experience goods to be less helpful than moderate reviews. For search goods, both extreme and moderate reviews can be helpful. An in-depth analysis of the text of reviewer comments, while beyond the scope of this paper, could yield additional insights.

Review Depth and Peer Comments

Review depth can increase information diagnosticity, and this is especially beneficial to the consumer if the information can be obtained without additional search costs (Johnson and Payne 1985). A reviewer's open-ended comments offer additional explanation and context to the numerical star ratings and can affect the perceived helpfulness of a review. When consumers are willing to read and compare open-ended comments from peers, the *amount* of information can matter. We expect the depth of information in the review content to improve diagnosticity and affect perceived helpfulness.

Consumers sometimes expend time and effort to evaluate alternatives, but then lack the confidence or motivation to make a purchase decision and the actual purchase. People are most confident in decisions when information is highly diagnostic. Tversky and Kahneman (1974) found that the increased availability of reasons for a decision increases the decision maker's confidence. Similarly, the arguments of senior managers were found to be more persuasive when they provided a larger quantity of information (Schwenk 1986). A consumer may have a positive inclination toward a product, but have not made the cognitive effort to identify the main reasons to choose a product, or to make a list of the pros and cons. Or, a consumer may be negatively predisposed toward a product, but not have the motivation to search and process information about other alternatives. In these situations, an in-depth review from someone who has already expended the effort is diagnostic, as it will help the consumer make the purchase decision.

The added depth of information can help the decision process by increasing the consumer's confidence in the decision. Longer reviews often include more product details, and more details about how and where the product was used in specific contexts. The quantity of peer comments can reduce product quality uncertainty, and allow the consumers to picture them-

selves buying and using the product. Both of these aspects can increase the diagnosticity of a review and facilitate the purchase decision process. Therefore, we hypothesize

> H2. *Review depth has a positive effect on the helpfulness of the review.*

However, the depth of the review may not be equally important for all purchase situations, and may differ depending on whether the consumer is considering a search good or an experience good. For experience goods, the social presence provided by comments can be important. According to social comparison theory (Festinger 1954), individuals have a drive to compare themselves to other people. Shoppers frequently look to other shoppers for social cues in a retail environment, as brand choice may be seen as making a statement about the individual's taste and values. Information that is personally delivered from a non-marketer has been shown to be especially credible (Herr et al. 1991).

Prior research has examined ways to increase the social presence of the seller to the buyer (Jiang and Benbasat 2004), especially as a way of mitigating uncertainty in the buyer–seller online relationships (Pavlou et al. 2007). Kumar and Benbasat (2006) found that the mere existence of reviews established social presence, and that online, open-ended peer comments can emulate the subjective and social norms of offline interpersonal interaction. The more comments and stories, the more cues for the subjective attributes related to personal taste.

However, reviews for experience products can be highly personal, and often contain tangential information idiosyncratic to the reviewer. This additional content is not uniformly helpful to the purchase decision. In contrast, customers purchasing search goods are more likely to seek factual information about the product's objective attributes and features. Since these reviews are often presented in a fact-based, sometimes bulleted format, search good reviews can be relatively short. The factual nature of search reviews implies that additional content in those reviews is more likely to contain important information about how the product is used and how it compares to alternatives. Therefore, we argue that while additional review content is helpful for all reviews, the incremental value of additional content in a search review is more likely to be helpful to the purchase decision than the incremental value of additional content for experience reviews. This leads us to hypothesize

> H3. *The product type moderates the effect of review depth on the helpfulness of the review. Review depth has a greater positive effect on the helpfulness of the review for search goods than for experience goods.*

To summarize, our model of customer review helpfulness (Figure 1) is an application of information economics theory and the paradigm of search versus experience goods. When consumers determine the helpfulness of a review, they take into account review extremity, review depth, and whether the product is a search good or an experience good.

Research Methodology ▬▬▬

Data Collection

We collected data for this study using the online reviews available through Amazon.com as of September 2006. Review data on Amazon.com is provided through the product's page, along with general product and price information that may include Amazon.com's own product review. We retrieved the pages containing all customer reviews for six products (see Table 1).

We chose these six products in the study based on several criteria. Our first criterion for selection was that the specific product had a relatively large number of product reviews compared with other products in that category. Secondly, we chose both search and experience goods, building on Nelson (1970, 1974). Although the categorization of search and experience goods continues to be relevant and widely accepted (Huang et al. 2009), for products outside of Nelson's original list of products, researchers have disagreed on their categorizations. The Internet has contributed to blurring of the lines between search and experience goods by allowing consumers to read about the experiences of others, and to compare and share information at a low cost (Klein 1998; Weathers et al. 2007). Given that products can be described as existing along a continuum from pure search to pure experience, we took care to avoid products that fell to close to the center and were therefore too difficult to classify.

We identify an experience good as one in which it is relatively difficult and costly to obtain information on product quality prior to interaction with the product; key attributes are subjective and difficult to compare, and there is a need to use one's senses to evaluate quality. We selected three goods that fit these qualifications well: a music CD, an MP3 player, and a video game. These are also typical of experience goods as classified in previous studies (see Bhattacharjee et al. 2006, Bragge and Storgårds 2007, Nelson 1970, Weathers et al. 2007). Purchase decisions on music are highly personal, based on issues more related to subjective taste than measurable attributes. It is difficult to judge the quality of a melody without hearing it. Even seeing the song's musical notes on a page would not be adequate for judging its overall quality.

An MP3 player has several objective, functional features such as storage capacity, size, and weight that can be judged prior to purchase. However, the MP3 player we chose (the iPod) is widely regarded as being popular more due its image and style than its functionality. Evaluation of the iPod rests heavily on interacting with the product and hearing its sound quality. A video game can also be described with some technical specifications, but the real test of quality is whether the game is entertaining and engaging. The entertainment quality is a subjective judgment that requires playing the game.

We define a search good as one in which it is relatively easy to obtain information on product quality prior to interaction with the product; key attributes are objective and easy to compare, and there is no strong need to use one's senses to evaluate quality. We found three goods that fit these qualifications: a digital camera, a cell phone, and a laser printer. Like our experience goods, these are representative of search goods used in previous research (see Bei et al. 2004; Nelson 1970; Weathers et al. 2007). The product descriptions on Amazon. com heavily emphasized functional features and benefits. Comparison tables and bullet points highlighted objective attributes. Digital cameras were compared on their image resolution (megapixels), display size, and level of optical zoom. Key cell phone attributes included hours of talk time, product dimensions, and network compatibility. Laser printers were compared on print resolution, print speed, and maximum sheet capacity. There is also an assumption that the products take some time to learn how to use, so a quick sampling or trial of the product is not perceived to be a good way of evaluating quality. Although using the product prior to purchase may be helpful, it is not as essential to assess the quality of the key attributes.

For each product, we obtained all of the posted reviews, for a total of 1,608 reviews. Each web page containing the set of reviews for a particular product was parsed to remove the HTML formatting from the text and then transformed into an XML file that separated the data into records (the review) and fields (the data in each review). We collected the following data:

(1) The star rating (1 to 5) the reviewer gave the product.
(2) The total number of people that voted in response to the question, "Was this review helpful to you (yes/no)?"
(3) The number of people who voted that the review was helpful.
(4) The word count of the review.

We excluded from the analysis reviews that did not have anyone vote whether the review was helpful or not. This led us to eliminate 21 reviews, or 1.3 percent of the total, resulting in a data set of 1,587 reviews of the 6 products.

Table 1. Products Used in the Study			
Product	**Description**	**Type**	**Sources**
MP3 player	5th generation iPod (Apple)	Experience	Weathers et al. 2007
Music CD	"Loose" by Nelly Furtado	Experience	Bhattacherjee et al. 2006 Nelson 1970 Weathers et al. 2007
PC video game	"The Sims: Nightlife" by Electronic Arts	Experience	Bragge and Storgårds 2007
Cell phone	RAZR V3 by Motorola	Search	Bei et al. 2004
Digital camera	PowerShot A620 from Canon	Search	Nelson 1970
Laser printer	HP1012 by Hewlett-Packard	Search	Weathers et al. 2007

Variables

We were able to operationalize the variables of our model using the Amazon data set. The dependent variable is helpfulness, measured by the percentage of people who found the review helpful (Helpfulness %). This was derived by dividing the number of people who voted that the review was helpful by the total votes in response to the "was this reviews helpful to you" question (Total Votes).

The explanatory variables are review extremity, review depth, and product type. Review extremity is measured as the star rating of the review (Rating). Review depth is measured by the number of words of the review (Word Count). Both of these measures are taken directly from the Amazon data for each review. Product type (Product Type) is coded as a binary variable, with a value of 0 for search goods and 1 for experience goods.

We included the total number of votes on each review's helpfulness (Total Votes) as a control variable. Since the dependent variable is a percentage, this could hide some potentially important information. For example, "5 out of 10 people found the review helpful" may have a different interpretation than "50 out of 100 people found the review helpful."

The dependent variable is a measure of helpfulness as obtained from Amazon.com. For each review, Amazon asks the question, "Was this review helpful?" with the option of responding "yes" or "no." We aggregated the dichotomous responses and calculated the proportion of "yes" votes to the total votes cast on helpfulness. The resulting dependent variable is a percentage limited to values from 0 to 100.

The descriptive statistics for the variables in the full data set are included in Table 2, and a comparison of the descriptive statistics for the search and experience goods subsamples are included in Table 3. The average review is positive, with an average star rating of 3.99. On average, about 63 percent of those who voted on a particular review's helpfulness found that review to be helpful in making a purchase decision. This indicates that although people tend to find the reviews helpful, a sizable number do not.

Analysis Method

We used Tobit regression to analyze the model due to the nature of our dependent variable (helpfulness) and the censored nature of the sample. The variable is bounded in its range because the response is limited at the extremes. Consumers may either vote the review helpful or unhelpful; there is no way to be more extreme in their assessment. For example, they cannot vote the review "essential" (better than helpful) or "damaging" (worse than unhelpful) to the purchase decision process. A second reason to use Tobit is the potential selection bias inherent in this type of sample. Amazon does not indicate the number of persons who read the review. They provided only the number of total votes on a review and how many of those voted the review was helpful. Since it is unlikely that all readers of the review voted on helpfulness, there is a potential selection problem. According to Kennedy (1994), if the probability of being included in the sample is correlated with an explanatory variable, the OLS and GLS estimates can be biased.

There are several reasons to believe these correlations may exist. First, people may be more inclined to vote on extreme reviews, since these are more likely to generate an opinion from the reader. Following similar reasoning, people may also be more likely to vote on reviews that are longer because the additional content has more potential to generate a reaction from the reader. Even the number of votes may be correlated with likelihood to vote due to a "bandwagon" effect.

Table 2. Descriptive Statistics for Full Sample

Variable	Mean	SD	N
Rating	3.99	1.33	1587
Word Count	186.63	206.43	1587
Total Votes	15.18	45.56	1587
Helpfulness %	63.17	32.32	1587

Table 3. Descriptive Statistics and Comparison of Means for Subsamples

Variable	Search (N = 559) Mean (SD)	Experience (N = 1028) Mean (SD)	p-value
Rating	3.86 (1.422)	4.06 (1.277)	**0.027†**
Word Count	173.16 (181.948)	193.95 (218.338)	**0.043‡**
Helpfulness Votes	18.05 (27.344)	13.61 (52.841)	0.064‡
Helpfulness %	68.06 (29.803)	60.52 (33.321)	**0.000‡**

†Using the Mann-Whitney test
‡Using the t-test

The censored nature of the sample and the potential selection problem indicate a limited dependent variable. Therefore, we used Tobit regression to analyze the data, and measured goodness of fit with the likelihood ratio and Efron's pseudo R-square value (Long 1997).

In H1, we hypothesized that product type moderates the effect of review extremity on the helpfulness of the review. We expect that for experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings. Therefore, we expect a nonlinear relationship between the rating and helpfulness, modeled by including the star rating as both a linear term (Rating) and a quadratic term (Rating[2]). We expect the linear term to be positive and the quadratic term to be negative, indicating an inverted U-shaped relationship, implying that extreme reviews will be less helpful than moderate reviews. Because we believe that the relationship between rating and helpfulness changes depending on the product type, we include interaction terms between rating and product type.

We include word count to test H2, that review depth has a positive effect on the helpfulness of the review. In H3, we expect that product type moderates the effect of review depth on the helpfulness of the review. To test H3, we include an interaction term for word count and product type. We expect that review depth has a greater positive effect on the helpfulness of the review for search goods than for experience goods.

The resulting model is[2]

$$Helpfulness \% = \beta_1 Rating + \beta_2 Rating^2 + \beta_3 Product\ type + \beta_4 Word\ Count + \beta_5 Total\ Votes + \beta_6 Rating \times Product\ type + \beta_7 Rating^2 \times Product\ type + \beta_8 Word\ Count \times Product\ type + \varepsilon$$

Results

The results of the regression analysis are included in Table 4. The analysis of the model indicates a good fit, with a highly significant likelihood ratio (p = 0.000), and an Efron's pseudo R-square value of 0.402.[3]

---

[2]We thank the anonymous reviewers for their suggestions for additional effects to include in our model. Specifically, it was suggested that we investigate the potential interaction of Rating and Word Count, and model the influence of Total Votes as a quadratic. When we included Rating × Word Count in our model, we found that it was not significant, nor did it meaningfully affect the level of significance or the direction of the parameter estimates. Total Votes[2] was significant (p < 0.0001), but it also did not affect the level of significance or the direction of the other parameter estimates. Therefore, we left those terms out of our final model.

[3]As a robustness check, we reran our analysis using an ordinary linear regression model. We found similar results. That is, the ordinary regression model did not meaningfully affect the level of significance or the direction of the parameter estimates.

Table 4. Regression Output for Full Sample				
Variable	**Coefficient**	**Standard Error**	**t-value**	**Sig.**
(Constant)	61.941	9.79	5.305	0.000
Rating	–6.704	7.166	–0.935	0.350
Rating²	2.126	1.118	1.901	0.057
Word Count	**0.067**	**0.010**	**6.936**	**0.000**
Product Type	**–45.626**	**12.506**	**–3.648**	**0.000**
Total Votes	–0.0375	0.022	–1.697	0.090
Rating × Product Type	**32.174**	**9.021**	**3.567**	**0.000**
Rating² × Product Type	**–5.057**	**1.400**	**–3.613**	**0.000**
Word Count × Product Type	**–0.024**	**0.011**	**–2.120**	**0.034**

Likelihood Ratio = 205.56 (p = 0.000, df 8, 1587

Efron's R² = 0.402

To test Hypothesis 1, we examined the interaction of rating and product type. Rating × Product type (p < 0.000) and Rating^2 × Product type (p < 0.000) were statistically significant. Product type moderates the effect of review extremity on the helpfulness of the review. To further examine this relationship, we split the data into two subsamples, search goods and experience goods. This is because in the presence of the interaction effects, the main effects are more difficult to interpret. The output from these two regressions are included in Tables 5 and 6.

For experience goods, there is a significant relationship between both Rating (p < 0.000) and Rating² (p < 0.001) and helpfulness. The positive coefficient for Rating and the negative coefficient for Rating² also indicates our hypothesized "inverted-U" relationship. For experience goods, reviews with extremely high or low star ratings are associated with lower levels of helpfulness than reviews with moderate star ratings. For search goods (Table 6), rating does not have a significant relationship with helpfulness, while Rating² does (p = 0.04). Therefore, we find support for H1. Product type moderates the effect of review extremity on the helpfulness of the review.

In H2, we hypothesize a positive effect of review depth on the helpfulness of the review. We find strong support for Hypothesis 2. For both search and experience products, review depth has a positive, significant effect on helpfulness. Word count is a highly significant (p < 0.000) predictor of helpfulness in both the experience good subsample (Table 5) and in the search good subsample (Table 6).

The results also provide strong support for H3, which hypothesizes that the product type moderates the effect of review

depth on the helpfulness of the review. This support is indicated by the significant interaction term Word Count × Product Type (p < 0.034) in the full model (Table 3). The negative coefficient for the interaction term indicates that review depth has a greater positive effect on the helpfulness of the review for search goods than for experience goods. A summary of the results of all the hypotheses tests are provided in Table 7.

Discussion

Two insights emerge from the results of this study. The first is that product type, specifically whether the product is a search or experience good, is important in understanding what makes a review helpful to consumers. We found that moderate reviews are more helpful than extreme reviews (whether they are strongly positive or negative) for experience goods, but not for search goods. Further, lengthier reviews generally increase the helpfulness of the review, but this effect is greater for search goods than experience goods.

As with any study, there are several limitations that present opportunities for future research. Although our sample of six consumer products is sufficiently diverse to support our findings, our findings are strictly generalizable only to those products. Future studies could sample a larger set of products in order to confirm that our results hold. For example, including different brands within the same product category would allow for an analysis of the potentially moderating effect of brand perception.

Second, the generalizability of our findings is limited to those consumers who rate reviews. We do not know whether those

Table 5. Regression Output for Experience Goods				
Variable	**Coefficient**	**Standard Error**	**t-value**	**Sig.**
(Constant)	5.633	8.486	0.664	0.507
Rating	**25.969**	**5.961**	**4.357**	**0.000**
Rating²	**−2.988**	**0.916**	**−3.253**	**0.001**
Word Count	**0.043**	**0.006**	**6.732**	**0.000**
Total Votes	−0.028	0.026	−1.096	0.273

Likelihood Ratio = 98.87 (p = 0.000, df 4, 1028)

Efron's R² = 0.361

Table 6. Regression Output for Search Goods				
Variable	**Coefficient**	**Standard Error**	**t-value**	**Sig.**
(Constant)	52.623	8.365	6.291	0.000
Rating	−6.040	6.088	−0.992	0.321
Rating²	**1.954**	**0.950**	**2.057**	**0.040**
Word Count	**0.067**	**0.008**	**8.044**	**0.000**
Total Votes	**−0.106**	**0.053**	**−2.006**	**0.045**

Likelihood Ratio = 98.87 (p = 0.000, df 4, 1028)

Efron's R² = 0.361

Table 7. Summary of Findings		
	Description	**Result**
H1	Product type moderates the effect of review extremity on the helpfulness of the review. For experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings.	Supported
H2	Review depth has a positive effect on the helpfulness of the review.	Supported
H3	The product type moderates the effect of review depth on the helpfulness of the review. Review depth has a greater positive effect on the helpfulness of the review for search goods than for experience goods.	Supported

reviews would be as helpful (or unhelpful) to those who do not vote on reviews at all. Future studies could survey a more general cross-section of consumers to determine if our findings remain consistent.

A third limitation is that our measures for review extremity (star rating) and review depth (word count) are quantitative surrogates and not direct measures of these constructs. Using data from the Amazon.com site has the advantage of being a more objective, data-driven approach than alternative approaches relying on subjective interpretations. For example,

subjectivity is required when determining whether comments or reviews are moderate, positive, or negative.

Still, qualitative analysis opens up several avenues for future research. One could analyze the text of the review and compare this to the star rating to determine how well the magnitude of the star rating matches with the review's content. In addition, one could use qualitative analysis to develop a more direct measure to create a more nuanced differentiation between moderate reviews and extreme reviews, as well as to develop a measure of review depth. For example, this type of

analysis could differentiate a three star review that contains conflicting but extreme statements from a three star review that contains multiple moderate statements.

Qualitative analysis could also be used to obtain additional quantitative data that can be incorporated into future studies. Pavlou and Dimoka (2006) performed a content analysis of comments regarding sellers on eBay and found that this yielded insights beyond what could be explained using purely quantitative measures. The analysis of the model indicates a good fit, with a highly significant likelihood ratio and a high Efron's pseudo R-square value, yet there are additional components of review helpfulness unaccounted for in this study. These data could be operationalized as new quantitative variables that could extend the regression model developed in this paper.

Finally, our regression model could also be extended to include other possible antecedents of review helpfulness, such as reviewer characteristics. This may be particularly relevant since review helpfulness is a subjective assessment, and could be influenced by the perceived credibility of the reviewer. Future studies could apply the search/experience paradigm to whether the reviewer's identity is disclosed (Forman et al. 2008) and the reviewer's status within the site (i.e., Amazon's "top reviewer" designations).

Conclusions

This study contributes to both theory and practice. By building on the foundation of the economics of information, we provide a theoretical framework to understand the context of online reviews. Through the application of the paradigm of search and experience goods (Nelson 1970), we offer a conceptualization of what contributes to the perceived helpfulness of an online review in the multistage consumer decision process. The type of product (search or experience good) affects information search and evaluation by consumers. We show that the type of product moderates the effect of review extremity and depth on the helpfulness of a review. We ground the commonly used measure of helpfulness in theory by linking it to the concept of information diagnosticity (Jiang and Benbasat 2004). As a result, our findings help extend the literature on information diagnosticity within the context of online reviews. We find that review extremity and review length have differing effects on the information diagnosticity of that review, depending on product type.

Specifically, this study provides new insights on the conflicting findings of previous research regarding extreme reviews. Overall, extremely negative reviews are viewed as less helpful than moderate reviews, but product type matters. For experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings, but this effect was not seen for search goods. Extreme reviews for experience products may be seen as less credible. Although we didn't specifically examine the relative helpfulness of negative versus positive reviews, future studies could address this question of asymmetry perceptions.

Our study provides an interesting contrast to the finding by Forman et al. (2008) that moderate book reviews are less helpful than extreme book reviews. In contrast, we found that for experience goods, reviews with extreme ratings are less helpful than reviews with moderate ratings, although this effect was not seen for search goods. Although books can be considered experience goods, they are a rather unique product category. Studies by Hu et al. (2008) and Li and Hitt (2008) look at the positive self-selection bias that occurs in early reviews for books. Our analysis of a wider range of experience and search goods indicates that additional insights can be gained by looking beyond one product type. Reviews and their effect on perceived helpfulness differ across product types.

We also found that length increases the diagnosticity of a search good review more than that of an experience good review. This is consistent with Nelson's (1970, 1974) classification of search and experience goods, in that it is easier to gather information on product quality for search goods prior to purchase. In the context of an online retailer, information comes in the form of a product review, and reviews of search goods lend themselves more easily to a textual description than do reviews of experience goods. For experience goods, sampling is required (Huang et al. 2009; Klein 1998). Additional length in the textual review cannot compensate or substitute for sampling.

This study is also has implications for practitioner audiences. Previous research has shown that the mere presence of customer reviews on a website can improve customer perception of the website (Kumar and Benbasat 2006). Sites such as Amazon.com elicit customer reviews for several reasons, such as to serve as a mechanism to increase site "stickiness," and to create an information product that can be sold to other online retailers. Reviews that are perceived as helpful to customers have greater potential value to companies, including increased sales (Chen et al. 2008; Chevalier and Mayzlin 2006; Clemons et al. 2006; Ghose and Ipeirotis 2006).

Our study builds on these findings by exploring the antecedents of perceived quality of online customer reviews. Our

findings can increase online retailers' understanding of the role online reviews play in the multiple stages of the consumer's purchase decision process. The results of this study can be used to develop guidelines for creating more valuable online reviews. For example, our results imply online retailers should consider different guidelines for customer feedback, depending whether that feedback is for a search good or an experience good. For a search good (such as a digital camera), customers could be encouraged to provide as much depth, or detail, as possible. For an experience good (such as a music CD), depth is important, but so is providing a moderate review. For these goods, customers should be encouraged to list both pros and cons for each product, as these reviews are the most helpful to that purchase decision. Reviewers can be incentivized to leave these moderate reviews. Currently, the "top reviewer" designation from Amazon is primarily determined as a function of helpfulness votes and the number of their contributions. Qualitative assessments could also be used, such as whether the reviews present pros and cons, in rewarding reviewers.

Our study also shows that online retailers need not always fear negative reviews of their products. For experience goods, extremely negative reviews are viewed as less helpful than moderate reviews. For search goods, extremely negative reviews are less helpful than moderate and positive reviews. Overall, this paper contributes to the literature by introducing a conceptualization of the helpfulness of online consumer reviews, and grounding helpfulness in the theory of information economics. In practice, helpfulness is often viewed as a simple "yes/no" choice, but our findings provide evidence that it is also dependent upon the type of product being evaluated. As customer review sites become more widely used, our findings imply that it is important to recognize that consumers shopping for search goods and experience goods may make different information-consumption choices.

Acknowledgments

References

Ba, S., and Pavlou, P. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly* (26:3), pp. 243-268.

Bakos, J. 1997. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science* (43:12), pp. 1676-1692.

Bei, L-T, Chen, E. Y. I., and Widdows, R. 2004. "Consumers' Online Information Search Behavior and the Phenomenon of Search and Experience Products," *Journal of Family and Economic Issues* (25:4), pp. 449-467.

Bhattacharjee, S., Gopal, R. D., Lertwachara, K., and Marsden, J. R. 2006. "Consumer Search and Retailer Strategies in the Presence of Online Music Sharing," *Journal of Management Information Systems* (23:1), pp. 129-159.

Bragge, J., and Storgårds, J. 2007. "Utilizing Text-Mining Tools to Enrich Traditional Literature Reviews. Case: Digital Games," in *Proceedings of the 30th Information Systems Research Seminar in Scandinavia*, Tampere, Finalnd, pp. 1-24.

Chen, P., Dhanasobhon, S., and Smith, M. 2008. "All Reviews Are Not Created Equal: The Disaggregate Impact of Reviews on Sales on Amazon.com," working paper, Carnegie Mellon University (available at SSRN: http://ssrn.com/abstract=918083).

Chen, P., Wu, S., and Yoon, J. 2004. "The Impact of Online Recommendations and Consumer Feedback on Sales," in *Proceedings of the 25th International Conference on Information Systems*, R. Agarwal, L. Kirsch, and J. I. DeGross (eds.), Washington, DC, December 12-14, pp. 711-724.

Chen, Y., and Xie, J. 2005. "Third-Party Product Review and Firm Marketing Strategy," *Marketing Science* (24:2), pp. 218-240.

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

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

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

Dabholkar, P. 2006. "Factors Influencing Consumer Choice of a 'Rating Web Site': An Experimental Investigation of an Online Interactive Decision Aid," *Journal of Marketing Theory and Practice* (14:4), pp. 259-273.

Eisend, M. 2006. "Two-Sided Advertising: A Meta-Analysis," *International Journal of Research in Marketing* (23), pp. 18-198.

Festinger, L. 1954. "A Theory of Social Comparison Processes," *Human Relations* (7), pp.117-140.

Ford, G. T., Smith, D. B., and Swasy, J. L. 1990. "Consumer Skepticism of Advertising Claims: Testing Hypotheses from Economics of Information," *Journal of Consumer Research* (16:4), pp. 433-441.

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

Ghose, A., and Ipeirotis, P. 2006. "Designing Ranking Systems for Consumer Reviews: The Impact of Review Subjectivity on Product Sales and Review Quality," in *Proceedings of the 16th Annual Workshop on Information Technology and Systems*

(available at http://pages.stern.nyu.edu/~panos/publications/wits2006.pdf).

Gretzel, U., and Fesenmaier, D. R. 2006. "Persuasion in Recommendation Systems," *International Journal of Electronic Commerce* (11:2), pp. 81-100.

Herr, P. M., Kardes, F. R., and Kim, J. 1991. "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective," *Journal of Consumer Research* (17:3), pp. 454-462.

Hu, N., Liu, L., and Zhang, J. 2008. "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects," *Information Technology Management* (9:3), pp. 201-214.

Huang, P., Lurie, N. H., and Mitra, S. 2009. "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing* (73:2), pp. 55-69.

Hunt, J. M., and Smith, M. F. "The Persuasive Impact of Two-Sided Selling Appeals for an Unknown Brand Name," *Journal of the Academy of Marketing Science* (15:1), 1987, pp. 11-18.

Jiang, Z., and Benbasat, I. 2004. "Virtual Product Experience: Effects of Visual and Functional Control of Products on Perceived Diagnosticity and Flow in Electronic Shopping," *Journal of Management Information Systems* (21:3), pp. 111-147.

Jiang, Z., and Benbasat, I. 2007. "Investigating the Influence of the Functional Mechanisms of Online Product Presentations," *Information Systems Research* (18:4), pp. 221-244.

Johnson, E., and Payne, J. 1985. "Effort and Accuracy in Choice," *Management Science* (31:4), pp. 395-415.

Kaplan, K. J. 1972. "On the Ambivalence-Indifference Problem in Attitude Theory and Measurement: A Suggested Modification of the Semantic Differential Technique," *Psychological Bulletin* (11), May, pp. 361-372.

Kennedy, P. 1994. *A Guide to Econometrics* (3rd ed.), Oxford, England: Blackwell Publishers.

Kempf, D. S., and Smith, R. E. 1998. "Consumer Processing of Product Trial and the Influence of Prior Advertising: A Structural Modeling Approach," *Journal of Marketing Research* (35), pp. 325-337.

Klein, L. 1998. "Evaluating the Potential of Interactive Media Through a New Lens: Search Versus Experience Goods," *Journal of Business Research* (41:3), pp. 195-203.

Kohli, R., Devaraj, S., and Mahmood, M. A. 2004. "Understanding Determinants of Online Consumer Satisfaction: A Decision Process Perspective," *Journal of Management Information Systems* (21:1), pp. 115-135.

Kotler, P., and Keller, K. L. 2005. *Marketing Management* (12th ed.), Upper Saddle River, NJ: Prentice-Hall.

Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K., and Camot, C. G. 1993. "Attitude Strength: One Construct or Many Related Constructs?," *Journal of Personality and Social Psychology* (65:6), pp. 1132-1151.

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

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

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks, CA: Sage Publications.

Mathwick, C., and Rigdon, E. 2004. "Play, Flow, and the Online Search Experience," *Journal of Consumer Research* (31:2), pp. 324-332.

Nelson, P. 1970. "Information and Consumer Behavior," *Journal of Political Economy* (78:20), pp. 311-329.

Nelson, P. 1974. "Advertising as Information," *Journal of Political Economy* (81:4), pp. 729-754.

Pavlou, P., and Dimoka, A. 2006. "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation," *Information Systems Research* (17:4), pp. 392-414.

Pavlou, P., and Gefen, D. 2004. "Building Effective Online Marketplaces with Institution-Based Trust," *Information Systems Research* (15:1), pp. 37-59.

Pavlou, P., and Fygenson, M. 2006. "Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior," *MIS Quarterly* (30:1), pp. 115-143.

Pavlou, P., Liang, H., and Xue, Y. 2007. "Uncertainty and Mitigating Uncertainty in Online Exchange Relationships: A Principal–Agent Perspective," *MIS Quarterly* (31:1), pp. 105-131.

Poston, R., and Speier, C. 2005. "Effective Use of Knowledge Management Systems: A Process Model of Content Ratings and Credibility Indicators," *MIS Quarterly* (29:2), pp. 221-244.

Presser, S., and Schuman, H. 1980. "The Measurement of a Middle Position in Attitude Surveys," *Public Opinion Quarterly* (44:1), pp. 70-85.

Schlosser, A. 2005. "Source Perceptions and the Persuasiveness of Internet Word-of-Mouth Communication," in *Advances in Consumer Research* (32), G. Menon and A. Rao (eds.), Duluth, MN: Association for Consumer Research, pp. 202-203.

Schwenk, C. 1986. "Information, Cognitive Biases, and Commitment to a Course of Action," *Academy of Management Review* (11:2), pp. 298-310.

Smith, D., Menon, S., and Sivakumar, K. 2005. "Online Peer and Editorial Recommendations, Trust, and Choice in Virtual Markets," *Journal of Interactive Marketing* (19:3), pp. 15-37.

Stigler, G. J. 1961. "The Economics of Information," *Journal of Political Economy* (69:3), pp. 213-225.

Todd, P., and Benbasat, I. 1992. "The Use of Information in Decision Making: An Experimental Investigation of the Impact of Computer-Based Decision Aids," *MIS Quarterly* (16:3), pp. 373-393.

Tversky, A., and Kahneman, D. 1974. "Judgment Under Uncertainty: Heuristics and Biases," *Science* (185:4157), pp. 1124-1131.

Weathers, D., Sharma, S., and Wood, S. L. 2007. "Effects of Online Communication Practices on Consumer Perceptions of Performance Uncertainty for Search and Experience Goods," *Journal of Retailing* (83:4), pp. 393-401.

About the Authors

Susan M. Mudambi is an associate professor of Marketing and Supply Chain Management, and an affiliated faculty member of the Department of Management Information Systems, in the Fox School of Business and Management at Temple University. Her research addresses strategic and policy issues in marketing, with special interests in the role of technology in marketing, international business, and business relationships. She has published more than a dozen articles, including articles in *Journal of Product Innovation Management, Industrial Marketing Management*, and *Management International Review*. She has a BA from Miami University, an MS from Cornell University and a Ph.D. from the University of Warwick (UK).

David Schuff is an associate professor of Management Information Systems in the Fox School of Business and Management at Temple University. He holds a BA in Economics from the University of Pittsburgh, an MBA from Villanova University, an MS in Information Management from Arizona State University, and a Ph.D. in Business Administration from Arizona State University. His research interests include the application of information visualization to decision support systems, data warehousing, and the assessment of total cost of ownership. His work has been published in *Decision Support Systems, Information & Management, Communications of the ACM*, and *Information Systems Journal*.

Appendix A

Differences in Reviews of Search Versus Experience Goods ▬▬▬▬▬▬

We observe that reviews with extreme ratings of experience goods often appear very subjective, sometimes go off on tangents, and can include sentiments that are unique or personal to the reviewer. Reviews with moderate ratings of experience goods have a more objective tone and reflect less idiosyncratic taste. In contrast, both extreme and moderate reviews of search goods often take an objective tone, refer to facts and measurable features, and discuss aspects of general concern. This leads us to our hypothesis (H1) that product type moderates the effect of review extremity on the helpfulness of the review. To demonstrate this, we have included four reviews from Amazon.com. We chose two of the product categories used in this study, one experience good (a music CD), and one search good (a digital camera). We selected one review with an extreme rating and one review with a moderate rating for each product.

The text of the reviews exemplifies the difference between moderate and extreme reviews for search and experience products. For example, the extreme review for the experience good takes a strongly taste-based tone ("the killer comeback R.E.M.'s long-suffering original fans have been hoping for…") while the moderate review uses more measured, objective language ("The album is by no means bad…But there are no classics here…"). The extreme review appears to be more of a personal reaction to the product than a careful consideration of its attributes.

For search goods, both reviews with extreme and moderate ratings refer to specific features of the product. The extreme review references product attributes ("battery life is excellent," "the flash is great"), as does the moderate review ("slow, slow, slow," "grainy images," "underpowered flash"). We learn information about the camera from both reviews, even though the reviewers have reached different conclusions about the product.

Experience Good: Music CD (REM's Accelerate)	
Excerpts from Extreme Review (5 stars)	***Excerpts from Moderate Review (3 stars)***
This is it. This really is the one: the killer comeback R.E.M.'s long-suffering original fans have been hoping for since the band detoured into electronic introspection in 1998. Peter Buck's guitars are front and centre, driving the tracks rather than decorating their edges. Mike Mills can finally be heard again on bass and backups. Stipe's vocals are as rich and complex and scathing as ever, but for the first time in a decade he sounds like he believes every word…It's exuberant, angry, joyous, wild - everything the last three albums, for all their deep and subtle rewards, were not. … Tight, rich and consummately professional, the immediate loose-and-live feel of "Accelerate" is deceptive. … Best of all, they sound like they're enjoying themselves again. And that joy is irresistible…	There's no doubt that R.E.M. were feeling the pressure to get back to being a rock band after their past three releases. …Peter Buck's guitar screams and shreds like it hasn't done in years and there is actually a drummer instead of a drum machine and looped beats…But you get the sense while listening to Accelerate, that Stipe and company were primarily concerned with rocking out and they let the songwriting take a back seat. The album is by no means bad. Several tracks are fast and furious as the title indicates. But there are no classics here, no songs that are going to return the band to the superstars they were in the 90's. And this album is not even close to their 80's output as some have suggested. … While Accelerate is a solid rock record, it still ranks near the bottom of the bands canon…

Search Good: Digital Camera (Canon SD1100IS)	
Excerpts from Extreme Review (5 stars)	***Excerpts from Moderate Review (3 stars)***
…Cannons are excellent cameras. The only reason i decided to replace my older cannon was because I am getting married, going to Hawaii and I wanted something a little newer/better for the trip of a lifetime. …Battery life is excellent. I had the camera for over a month, used it a lot (especially playing around with the new features) and finally just had to charge the battery… The flash is great- …I only had red eyes in about half the shots (which for me is great)…There are a ton of different options as to how you can take your photo, (indoor, outdoor, beach, sunrise, color swap, fireworks, pets, kids etc etc)… Cannons never disappoint in my experience!	Just bought the SD1100IS to replace a 1-1/2 year old Casio Exilim EX-Z850 that I broke. … But after receiving the camera and using it for a few weeks I am beginning to have my doubts. ... 1. Slow, slow, slow - the setup time for every shot, particularly indoor shots, is really annoying…2. Grainy images on screen for any nightime indoor shots. 3. Severely under-powered flash - I thought I read some reviews that gave it an OK rating at 10 feet...try about 8 feet and you might be more accurate. Many flash shots were severely darkened by lack of light...and often the flash only covered a portion of the image leaving faces dark and everything else in the photo bright…