Decision Support Systems 79 (2015) 12–23

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Can online product reviews be more helpful? Examining characteristics of information content by product type

Danny Weathers [a,1,2], Scott D. Swain [a,1,2], Varun Grover [b,*]

[a] Department of Marketing, Clemson University, United States
[b] Department of Management, Clemson University, United States

ARTICLE INFO

Article history:
Received 28 August 2014
Received in revised form 21 May 2015
Accepted 20 July 2015
Available online 2 August 2015

Keywords:
Product review
Search-experience goods
Helpfulness
Information uncertainty
Information equivocality
Word-of-mouth

ABSTRACT

Many online retailers and other product-oriented websites allow people to post product reviews for use by shoppers. While research indicates that these reviews influence consumers' shopping attitudes and behaviors, questions remain about how consumers evaluate the product reviews themselves. With the current research, we introduce a new methodology for identifying the review factors that shoppers use to evaluate review helpfulness, and we integrate prior literature to provide a framework that explains how these factors reflect readers' general concerns about the diagnosticity (uncertainty and equivocality) and credibility (trust and expertise) of electronic word-of-mouth. Based on this framework, we offer predictions about how the relative importance of diagnosticity and credibility should vary systematically across search and experience product types. By analyzing secondary data consisting of over 8000 helpfulness ratings from product reviews posted by shoppers on Amazon.com, we find that, while review content affects helpfulness in complex ways, these effects are well explained by the proposed framework. Interestingly, the data suggest that review writers who explicitly attempt to enhance review diagnosticity or credibility are often ineffective or systematically unhelpful. Our findings have implications for both IS developers and retailers for designing online decision support systems to optimize communication practices and better manage consumer-generated content and interactions among consumers.

1. Introduction

Between non-retail websites dedicated to eliciting and aggregating consumer feedback (e.g., Epinions.com, Rateitall.com, Yelp.com) and the many online retailers and manufacturers who have followed suit (e.g., Amazon.com, Sears.com, Dell, Levi's), online shoppers have increasingly greater access to other shoppers' opinions and reviews of products. The availability of consumer product reviews, henceforth referred to as reviews, is likely to continue proliferating for at least two reasons. First, the daunting number of online options leads consumers to value reviews both as a filtering mechanism and as an important source of information about product characteristics often difficult to assess in an online environment [78,94]. Indeed, due to the evolution of online retailing and social media, consumers now enter the marketplace with the expectation to access reviews [74,81]. Second, a desire to increase throughput and reduce costs associated with customer support and product returns [86] is likely to motivate retailers to facilitate reviews in efforts to enhance customer involvement with the website and improve consumers' decision-making.

Recent research in the areas of information systems and marketing provides a number of insights on reviews. In terms of antecedents, studies identify the characteristics and motivations of those who write reviews [18,35,71], including strategic fake reviews written on behalf of organizations [57] and non-strategic deceptive reviews written by individual, non-purchasers [1]. Complementing this perspective, another stream of research examines the outcomes of reviews, finding that reviews affect aggregate consumer behavior as reflected in sales, profits, and viewership [17,21,25,31,54,76,93], website and product evaluations [12,39,49,72], competitive intelligence [90], and individual consumer choice [24,28,34].

How consumers decide whether they can rely on a particular review is less examined. Speaking to the applied importance of this issue, many retail websites elicit, summarize, and publish consumers' feedback on the "usefulness" or "helpfulness" of individual reviews. Despite recent studies examining the impact of a number of review attributes or characteristics on consumer perceptions of helpfulness, a cohesive framework of review helpfulness has yet to emerge. Given the prevalence of reviews, as well as the strategic importance of managing this information, such a framework should prove useful for both researchers and practitioners by revealing factors and contingencies that enhance or detract from review helpfulness.

* Corresponding author. Tel.: +1 864 656 3773.
E-mail addresses: pweath2@clemson.edu (D. Weathers), sdswain@clemson.edu (S.D. Swain), vgrover@clemson.edu (V. Grover).
[1] The first two authors contributed equally to this article.
[2] Tel.: +1 864 656 5963.

http://dx.doi.org/10.1016/j.dss.2015.07.009
0167-9236/

Integrating concepts and findings from prior research, we propose that perceptions of helpfulness can be understood in terms of the interactive effects of three factors: review credibility, review diagnosticity, and product type. Going beyond extant work, we distinguish between aspects of credibility and diagnosticity based on principles of consumer information search. Whereas credibility is conceptualized as a function of trust (i.e., providing unbiased information) and expertise (i.e., being competent to provide the information), diagnosticity is conceptualized in terms of consumers' desire to not only reduce uncertainty (i.e., a lack of information), but also reduce equivocality (i.e., the plausibility of multiple, conflicting interpretations). Additionally, we predict that the type of product being reviewed (search versus experience goods) systematically impacts the influence of credibility and diagnosticity on review helpfulness. Finally, we draw on recent research on the social psychology of hubris and egocentrism to make novel predictions about the conditions in which consumers are likely to reject review writers' signals of diagnosticity and credibility. Fig. 1 illustrates the general framework.

We employ two methodological refinements with respect to existing research to enhance confidence in the validity of the findings. First, extant work relies on researcher judgment to select review characteristics for study. While this is a reasonable strategy given the linguistic complexity and contextualization of reviews, we maintain for this same reason that it is also critical to corroborate and refine these judgments in consultation with consumers to ensure that selected review characteristics are psychologically meaningful. We present a method for doing so. Second, consistent with existing studies, we leverage a search/experience product classification framework to understand how review characteristics may operate differentially across product categories. However, rather than basing product classification solely on researcher judgment, we introduce a multi-dimensional scaling approach that integrates researcher *and* consumer judgments. This approach recognizes that consumers' intuitive product classifications are sensitive to variations in factors such as experience, usage occasions, and retail channel characteristics [4,59,83].

We assess the validity of the proposed framework of review helpfulness using two pretests and a main study. Pretest 1 validates the classification of a set of commonly purchased products. Pretest 2 identifies factors that are prominent in consumers' assessments of review helpfulness, thus focusing our efforts. In the main study, we analyze 8327 helpfulness ratings for reviews posted by shoppers on Amazon.com. We find broad support that review helpfulness depends not only on the ability to reduce decision uncertainty, but also on reducing informational equivocality while conveying trustworthiness and expertise. These effects vary by product type in ways that are predicted by the proposed framework. We discuss the implications of the findings for managing consumer-generated web content through online support systems and for research on IS mediated consumer-to-consumer communication.

2. Literature review and hypotheses

In this section, we describe the components of our framework: diagnosticity, credibility, and product type. We then present the review content factors on which we focus, and we develop predictions regarding the effects of these factors on helpfulness.

2.1. Overview of diagnosticity and credibility

We argue that the primary reason consumers read reviews is to move toward the ultimate goal of making a purchase decision (i.e., buy/no buy). Reducing uncertainty about the product should help achieve this goal and, indeed, reviews are seen as diagnostic to the extent that they reduce product uncertainty (e.g., [61]). However, the information search and word-of-mouth literatures suggest that reviews may play another important function. Specifically, reviews may not only reduce uncertainty but also reduce equivocality. Whereas uncertainty refers to a lack of information, equivocality refers to ambiguity or the plausibility of multiple interpretations [15,16]. Consumers are motivated to reduce equivocality because high levels of equivocality adversely affect decision-making (e.g., [53,85]).

However, product-related information alone is unlikely to fully determine consumers' perceptions of review helpfulness, as this implies that consumers take all reviews at face value. Rather, as research has found, consumers are also likely to consider the source of the information. Product reviews are mediated word-of-mouth, a form of interpersonal communication in which neither participant is a marketing source [8]. A critical factor determining the influence of word-of-mouth information is the perceived credibility of the source [6,63], particularly in the context of web-based commerce [13,26,92]. There are two dimensions of credibility: expertise, or the extent to which the communicator is perceived as a source of valid assertions (i.e., competent), and trustworthiness, or the extent to which the communicator is perceived as a source of unbiased assertions [38].

2.2. Product type: search versus experience goods

While several product classification paradigms are potentially relevant for understanding review helpfulness, the search/experience paradigm has proven particularly useful for explaining online shopping behavior (e.g., [41,70]) and for understanding consumer evaluations of online product reviews (e.g., [4,42,61,64,87]). Nelson [62] distinguishes between search and experience goods based on the extent to which shoppers *can* experience the goods prior to purchase. Others argue that because this search/experience distinction can vary across retail channels for a given product, a classification paradigm that is less channel-dependent is beneficial [59,83]. Consistent with this latter

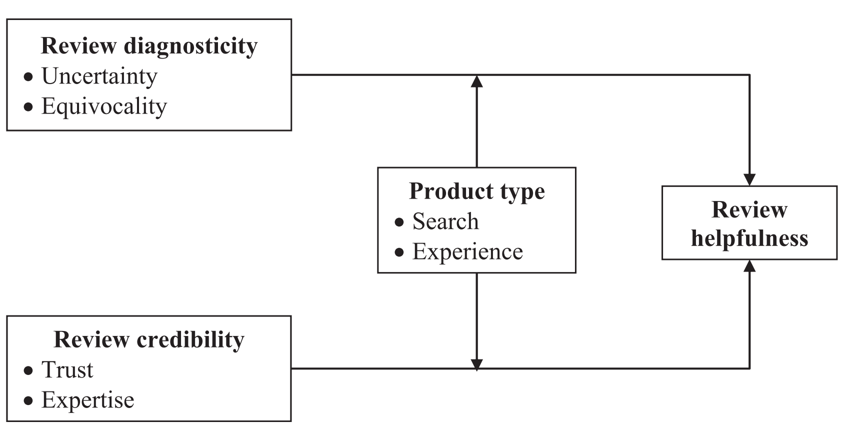

Fig. 1. Antecedents of perceived review helpfulness.

perspective, we conceptualize the search/experience distinction as the extent to which shoppers *feel the need* to directly experience goods to evaluate their quality. The greater (less) the perceived need to directly experience a product, the more experience (search) qualities the product possesses. We view goods as a bundle of attributes, and we consider classification at the good, rather than attribute, level.

2.3. Characterizing the content of consumer product reviews

Reviews differ in a number of ways that can be connected to the abstract concepts of diagnosticity and credibility. While research to date focuses on a subset of review content factors judged by *researchers* as potentially determinant of helpfulness, the *consumer* insights of Pretest 2, subsequently presented in detail, guide our efforts. Accordingly, we triangulate our data with existing evidence by reconsidering three factors that have previously been shown to impact helpfulness: *claims of expertise*, or whether the reviewer makes explicit claims to be an expert (e.g., [52,87]); *review valence*, or whether the review provides a positive or negative evaluation of the product (e.g., [71, 87]); and *review balance*, or whether the review provides both positive and negative information about the product (e.g., [69,87,88]).

More focally, we consider four factors that have not yet been directly examined in the literature. *Listing of features* refers to whether the reviewer explicitly lists product features without evaluating these features, and *descriptions of usage situations* refers to whether the reviewer provides information on specific usage or consumption experiences. These factors can be tied to research which reveals that concrete (i.e., objective) versus abstract (i.e., subjective) claims impact review helpfulness [52]. Listing features is a concrete claim, while describing usage situations is more abstract. *References to other reviews* refers to whether the reviewer explicitly mentions another review, and *comparisons to other brands* refers to whether the reviewer explicitly compares the product being reviewed to another brand of the same product type.

Finally, we include three control variables. *Review extremity*, or the extent to which a review is atypically positive or negative, can impact helpfulness, though the effect may depend on the type of product being reviewed (e.g., [61]). *Review length*, or the number of words in the review, has been shown to affect helpfulness (e.g., [61,87]), though length is also likely to be a function of the factors mentioned earlier. *Product price* is the price of the product reviewed. Consistent with our framework, we now argue that each focal factor relates to consumer perceptions of helpfulness either through impacting uncertainty, equivocality, reviewer expertise, or reviewer trustworthiness, and that the search-experience dichotomy moderates each of these effects.

2.4. Hypotheses

As discussed, we focus on review characteristics that should contribute to the consumer goals of reducing uncertainty and equivocality and also provide information on reviewer expertise and trustworthiness. We now develop predictions for the effects of these characteristics on perceptions of review helpfulness. In each hypothesis, we provide the variable names associated with the review characteristics in the main study.

2.4.1. Review balance

Whether the reviewer is thought to provide a complete accounting of the product should affect consumers' perceptions of the reviewer's trustworthiness. As such, balanced reviews containing negative and positive evaluations should enhance trustworthiness compared to one-sided reviews containing only negative or only positive evaluations [87]. By indicating potential benefits and problems, balanced (versus one-sided) reviews suggest to consumers that they have received a more honest picture of the product. This principle is well-supported in the domain of advertising where consumers perceive two-sided communications as more credible than one-sided communications (e.g., [43]). However, because negative evaluations of experience attributes tend to enhance source credibility more than negative evaluations of search attributes [65], the effect of balance should vary by product type. Thus, we predict:

H1. Providing a balanced review with both positive *and* negative information (BALANCE) has a positive effect on helpfulness for experience goods. This effect is attenuated for search goods.

2.4.2. Reviewer claims of expertise

Reviewers often try to convey their expertise. Ostensibly this is a sensible strategy. Consumers tend to view experts as topical authorities and, by extension, as valid sources of information about related products (e.g., [77,95]). Empirical studies also find that reviews are more impactful when they come from expert sources [58,74] and that reviews garner higher ratings when reviewers provide identity-descriptive information [25]. Consequently, indications of reviewer expertise should enhance the helpfulness of a review [23]. However, unlike previous research, we argue that the *manner* in which a reviewer conveys expertise is critical to determining the impact on helpfulness.[3]

Sometimes reviewers convey expertise in an indirect and noncomparative manner (e.g., “I owned my previous grill for 10 years and grilled out 3–4 times a week” or “I talked with my doctor, and he recommended this [skin care] product for the following reasons …”). In contrast, other reviewers assert expertise through self-superiority claims or blatant social comparisons (e.g., “I have won my cul-de-sac grilling championship for the past 3 years running …”). Research on social perceptions of prideful self-enhancement suggests that people are significantly more averse to the latter form of expression, not because it violates social norms but because it implies that the reviewer is arrogant or holds a negative view of others [37,79].

In the context of product reviews, reviewers who state the credentials of others (e.g., “this information came from my doctor”) are not likely to be viewed as prideful because they are not commenting on their own abilities. Similarly, reviewers who simply indicate the extent of their personal experience (e.g., “I owned my previous grill for 10 years and grilled out 3–4 times a week”) avoid the implication of self-superiority and instead communicate expertise in a noncomparative manner. These manners of conveying expertise should enhance perceptions of review helpfulness. However, reviewers who directly assert their expert credentials risk conveying condescension, as if they are saying “I'm more of an expert than you are so just listen to my opinion.” Readers are likely to find this sort of reviewer less likable and, consequently, less credible. Thus, the review should be rated as less helpful.

We expect the nature of the product to moderate the effect of expertise. Specifically, since online shoppers can largely evaluate search goods (but not experience goods) based on information provided by retailers, a reviewer's expertise is likely to be less important. Thus, the enhancing effects of claims of expertise based on direct personal experience or the experience of others will tend to be attenuated for search goods. No such attenuation is expected for claims of expertise that imply a negative view of others because consumers' aversive reactions to the review focus on the likeability of the reviewer rather than the content of the review. Thus:

H2(a). Claiming expertise based on direct experience or the experience of others (EXPERT1) has a positive effect on helpfulness for experience goods. This effect is attenuated for search goods.

[3] We thank an anonymous reviewer for this suggestion.

H2(b). Claiming expertise based on credentials of the reviewer (EXPERT2) has no or adverse effects on helpfulness for both experience and search goods.

2.4.3. References to other brands

A reference to another brand is a diagnostic cue for categorizing information as "helpful" or "unhelpful" [73]. Explicitly referencing other brands suggests that the reviewer has experience with the product category and, therefore, is qualified to evaluate the product. Further, comparing a product to another brand or brands should facilitate the end goal of making a choice. These comparisons help shoppers rank order options and rule out alternatives, regardless of whether the product being reviewed dominates, or is dominated by, the referenced brand(s). However, the effect of referring to other brands is likely to differ by product type. Brand names tend to be more important to consumers for experience goods than for search goods due to search goods' tangible attributes and generally lower perceived risk [22]. For search goods, risk is often sufficiently minimized through written marketing communications [10], and consumers are less concerned about making a mistake when choosing a brand [5]. However, when conveying the essence of products with written or quantifiable descriptions is difficult, as with experience goods, brand names are more important to consumers [30]. Thus:

H3. Referring to other brands (OTHERBRANDS) has a positive effect on helpfulness for experience goods. This effect is attenuated for search goods.

2.4.4. References to other reviews

A salient goal for consumers who post reviews is to appear thoughtful and helpful [18]. Reviewers may pursue this goal by referring to other reviews, potentially providing richer information for readers through integrating multiple perspectives. Reviewers can refer to other reviews in three ways. First, reviewers can reconcile multiple conflicting reviews. A reviewer of a skin care product stated, "Some of you have reported good results, and others have reported bad results. The reason for this is because the product's effectiveness depends on your skin type." Second, reviewers can agree with the comments of others. A reviewer of a laptop computer stated, "It overheats, which other reviewers have mentioned." Third, reviewers can disagree with the comments of others. A reviewer of a DVD player stated, "Unlike other reviewers, I found the remote works great." The first two types of references should reduce equivocality by building consensus about the product. However, the third reference type may have the unintended consequence of increasing equivocality because readers will often be left unsure as to which claim is more accurate or which reviewer is more credible [14,60].[4]

If multiple people agree about the performance of a product or if a reviewer can explain why the product works for some people but not others, the reader of the review should feel more confident in gauging how the product will perform for her/him. Given the idiosyncratic nature of experience goods, such consensus building or reconciliation is likely to be more critical for experience goods than for search goods [93]. Thus:

H4(a). Referring to other reviews by agreeing with the reviews or integrating conflicting reviews (OTHERREVIEWS1) has a positive effect on helpfulness for experience goods. This effect is attenuated for search goods.

H4(b). Referring to other reviews by disagreeing with the reviews (OTHERREVIEWS2) has a negative effect on helpfulness for both experience and search goods.

[4] We thank an anonymous reviewer for this insight.

2.4.5. Description of usage situations

While reviews may contain simple, evaluative product judgments (e.g., "This camera is a good value."), they may also provide information about how the reviewer uses the product (e.g., "I've used the camera in low light to photograph the night sky."). Such statements can influence the formation of consumption visions, or "visual images of certain product-related behaviors and their consequences [leading to] concrete and vivid mental images that enable consumers to vicariously experience the self-relevant consequences of product use" ([82], pp. 27, 31). Consumption visions provide consumers with clarity about "how product attributes relate to the self via the consequences of product use" ([66], p. 283). Thus, descriptions of product usage should reduce equivocality and, in so doing, enhance the perceived helpfulness of the review. However, because contact with a product is more important for evaluating experience goods than search goods, descriptions of product usage which facilitate consumption visions are likely to have a greater impact on reducing equivocality about experience goods than search goods [83]. Thus:

H5. Descriptions of usage situations (USAGE) have a positive effect on helpfulness for experience goods. This effect is attenuated for search goods.

2.4.6. Listing features

A consumer who has written a review is generally further along in the decision-making process than a consumer who seeks reviews. This creates a potential mismatch in the type or level of information sought and type or level of information provided. Consumers tend to use simpler and less cognitively effortful decision rules in earlier stages than in later stages of the decision-making process [27]. For example, early in the process, consumers may want to know whether a product possesses certain features (i.e., a simple "yes" or "no," which can be determined if the feature is listed by the reviewer) as opposed to how well those features function (i.e., a more complex evaluative judgment). Thus, reviews that list product features may (1) close information gaps for readers by disclosing product information that the shopper missed when provided by the retailer or (2) remind readers about, or focus readers on, particularly determinant features. Consequently, consumers should perceive reviews as more helpful when they list product features than when they do not, as shopper uncertainty is reduced. Consistent with our feature-based conceptual distinction between search and experience goods, this effect should be stronger for search goods. Thus:

H6. Listing product features (FEATURES) has a positive effect on helpfulness for search goods. This effect is attenuated for experience goods.

3. Methodology

Prior to testing the hypotheses, we present two pretests designed to improve the validity of the constructs in the main study. First, the framework predictions involve a theoretical distinction between search and experience goods. In Pretest 1, we introduce a method for using consumer input to verify this distinction. Second, extant research relies on researcher judgment to identify which review characteristics might impact perceptions of helpfulness. In Pretest 2, we introduce an approach for allowing consumers to help identify which review characteristics are relevant for focal products.

3.1. Pretest 1 — validation of the search/experience product classification

3.1.1. Procedure and measures

We selected eight product categories for study, tentatively judging four of the categories to possess relatively more experience qualities (vacuum cleaners, outdoor grills, skin care products, music CDs) and four to possess relatively more search qualities (DVD players, laptop computers, digital camcorders, books). To verify these judgments, we administered a multidimensional scaling task to ascertain whether, unprompted, consumers think in search/experience terms when considering product evaluability in an online environment. We then examined the extent to which the dimensional scores correlated with a direct measure of consumers' perceptions of the search/experience product qualities.

In the multidimensional scaling task, we asked respondents to sort the eight product categories into three groups based on "how well you think you could evaluate the products if you were shopping for them on the Internet." The three groups were "best able to evaluate," "moderately able to evaluate," and "least able to evaluate." Thus, respondents perceived products within each group as more similar to each other than to product(s) in the remaining group(s). Respondents could place any number of products in each group. The number of times respondents grouped each pair of products together served as a measure of similarity in a proximity matrix. We performed non-metric multidimensional scaling on this matrix. In the subsequent direct measurement task, respondents indicated how important they felt it was to experience each product "in person" for the purpose of evaluating it (1 = not at all important, 10 = very important). Higher values indicate greater experience (or lower search) qualities.

3.1.2. Results

We performed multidimensional scaling on classification data provided by sixty-nine MBA students (average age = 24.2, 46% female). The scaling procedure treated the data as ordinal and broke ties per the primary approach [9]. We implemented the SMACOF algorithm (Scaling by MAjorizing a Complicated Function) to minimize the model loss function (i.e., stress) in a non-metric multidimensional scaling analysis [20]. We did not impose any external constraints, and we employed a multi-start procedure (100 random starts and 1000 maximum iterations per start) to guard against obtaining local optima.

Our initial analysis revealed strong support for a one-dimension solution. However, the book and music CD scored close to zero on this dimension, indicating that respondents did not strongly perceive these products as falling on either side of this dimension. We dropped these products and re-ran the analysis with the remaining six products. The final product stimulus configuration explained 99.3% of the original dispersion in the transformed proximities. Fig. 2(a) illustrates this goodness of fit; the fitted distances between products (y-axis) closely mirror the empirical proximities collected from the respondents (x-axis). The diagonal line indicates the location of points under perfect fit. Thus, a single dimension provides a good representation of respondents' judgments of product evaluability with respect to online shopping, and, by implication, additional dimensions are not necessary to account for the respondents' judgments.

We observe that the relative positions of the product stimuli on the derived dimension are consistent with a search/experience configuration. Specifically, Fig. 2(b) shows that the DVD player, laptop computer, and digital camcorder (selected as search goods) are located on one end of the dimension and the vacuum cleaner, outdoor grill, and skin care products (selected as experience goods) are located on the other end.

The direct measures of respondents' perceptions of the search/experience nature of the products converge with the indirect measure provided by multidimensional scaling. The means from the direct measure correlate highly with the multidimensional scaling coordinates ($r = .78, p = .03$). Further, the mean value of the direct measure for the DVD player/laptop computer/digital camcorder group is significantly lower than the mean value for the vacuum cleaner/outdoor grill/skin care product group ($M = 4.94$ versus $M = 6.41$, $t = -4.36$, $p < .01$).

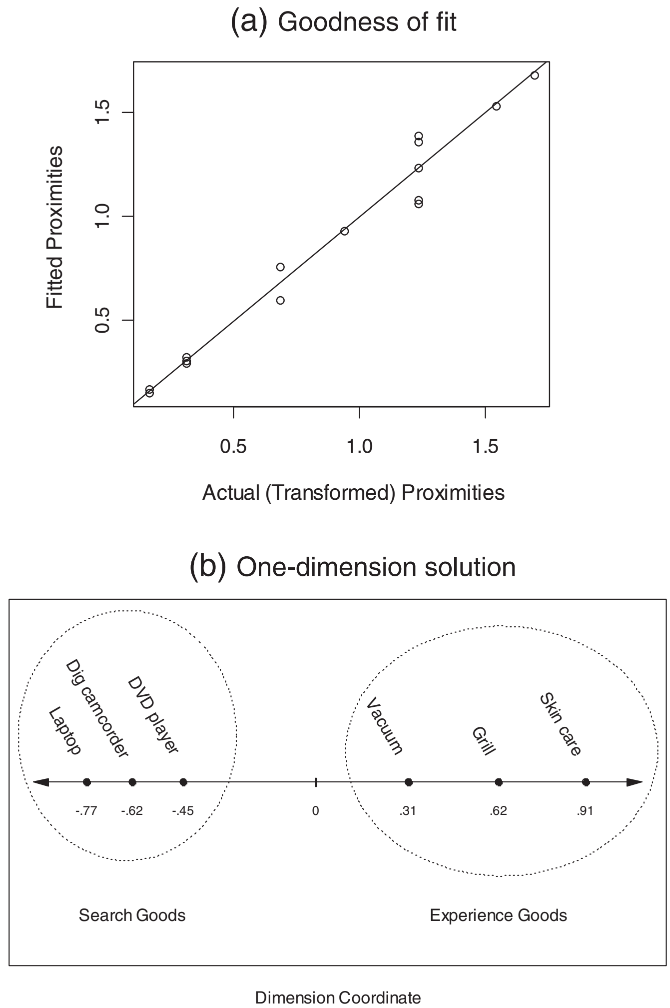

Fig. 2. Pretest 1 results.

3.2. Pretest 2 — identifying relevant review characteristics

While research to date explores various review characteristics that may impact helpfulness, to our knowledge, no prior work obtains direct insight from consumers as to what they themselves look for in a helpful review. Pretest 2 does so, thus guiding the selection of the review factors on which we based the hypotheses.

3.2.1. Procedure and sample

Pretest 1 identified three products as search goods (DVD players, laptop computers, and digital camcorders) and three products as experience goods (vacuum cleaners, outdoor grills, and skin care products). For consistency across studies, we use these same six products in Pretest 2 as well as the main study. As part of a class project on collecting and analyzing open-ended data, students enrolled in a research course were each assigned a randomly selected subset of three of the six products. They were instructed to visit Amazon.com to find any one product review for each of the three product categories. They then emailed the three reviews to five people, with instructions for the recipients to read the reviews and list "all the thoughts you had in deciding whether the review would be helpful or unhelpful to you in shopping for the product." Responses were returned to the students and compiled by the authors. The usable sample consisted of 251 respondents, with a mix of college students and non-student adults, who provided thoughts on 753 reviews. The responses contained a total of 2138 helpfulness-related thoughts (an average of 2.83 thoughts per review).

Trained research assistants coded the responses. The procedure was "triple blind" in the sense that the research assistants, data collectors, and respondents were unaware of any hypotheses.

3.2.2. Results

We allowed response categories to emerge from the data. For each product, we determined the percentage of respondents who provided a comment that fell in each category, and we averaged these percentages across products. On average, 79.7% of respondents made comments that reflected evaluations, not characteristics, of reviews. For example, respondents commented that reviews were "informative," "convincing," "lacked important information," "would save me time," "contained irrelevant information," or "would convince me to buy the product." Such comments support our contention that shoppers read reviews with an eye toward reducing uncertainty or equivocality and gauging the source of the review. For example, a review that "lacked important information" would not reduce uncertainty.

In terms of specific review characteristics, an average of 41.6% of respondents mentioned issues related to reviewer expertise or trustworthiness, including comments such as "I like that the reviewer was an expert" or "The reviewer was a 40-year-old mother, so I couldn't relate to her." An average of 22.6% of respondents referred to usage situations (e.g., "I like how it discussed how to use a memory card"), and comments about product features were made by an average of 22.3% of respondents (e.g., "It was helpful because it talked about the battery life"). On average, 13.2% of respondents cited references to other brands (e.g., "It was helpful to compare the computer to a Dell"), and 12.5% of respondents mentioned the presence of both positive and negative evaluations (e.g., "I like that the review listed pros and cons of the product"). An average of 7.7% of respondents made references to valenced information (e.g., "I like that the reviewer talked about positive aspects of the product"). An average of 1.9% of respondents mentioned references to other reviews (e.g., "The reviewer corrected mistakes by other reviewers"). An average of 9.8% of respondents cited review length (e.g., "The review was too long"). We ultimately use length as a control variable.[5]

The qualitative findings of Pretest 2 are not indicative of the importance of various review characteristics for judging helpfulness. The percentage of respondents commenting on a particular characteristic was largely a function of the content of the reviews with which they were provided. For example, a respondent may not have mentioned references to other brands, not because such references are unimportant, but because the review he/she evaluated did not refer to other brands. Overall, the results of Pretest 2 guided the selection of the review characteristics on which the hypotheses are based by indicating that the characteristics are relevant to consumers for the task of judging helpfulness.

3.3. Main study

3.3.1. Procedure

We tested the hypotheses using secondary data consisting of product reviews collected from Amazon.com. On Amazon, people can post reviews of products, and shoppers can indicate whether or not the reviews are helpful by responding to the question "Was this review helpful to you? Yes/No." Within each of the six product categories, we randomly selected 30 reviews, subject to the reliability constraint that at least 10 people had indicated whether or not the review was helpful. The final sample contained 180 product reviews (90 search goods, 90 experience goods) with 8327 helpfulness ratings (an average of 46.3 ratings per review).

Table 1 describes the variables measured for each review, along with the mapping of each variable onto our theoretical framework. Following the procedure advocated by Srnka and Koeszegi [75], two research assistants blind to the purpose of the study independently coded the reviews. Inter-rater agreement was 85.8%, and Cohen's κ was .702, indicating high inter-rater reliability [51]. The research assistants resolved coding discrepancies through discussion. Table 2 provides summary statistics for the variables listed in Table 1, both for the overall sample of reviews and for the reviews for each product category.

The review was the unit of coding and analysis. To illustrate the variable coding described in Table 1, we present examples from reviews contained in our sample. For a review of a digital camcorder, 72 of 74 people found the review to be helpful (HELP). The price of the camcorder was \$629.99 (PRICE). The review contained 316 words (LENGTH). In total, 48 people reviewed this camcorder, and the average evaluation was 4.5 stars (OVERALLEVAL). This particular reviewer gave the product 5 stars (INDIVIDUALEVAL), thus EVALDIFF, which captures review extremity, is $|5.0 - 4.5| = 0.5$. In a review for a DVD player, the reviewer stated "It has great picture and sound quality," thus POSINFO is coded "1." The reviewer also stated "A con is that it is slow to load disks," thus NEGINFO is coded "1." Because the reviewer provided both positive and negative comments about the product's features, BALANCE is coded "1." A reviewer of a vacuum cleaner stated, "I bought this vacuum to replace a vacuum that I had used almost daily for 7–8 years," thus EXPERT1 is coded "1." A reviewer of a grill stated, "I have won my cul-de-sac grilling championship for the past 3 years running," thus EXPERT2 is coded "1." In a vacuum cleaner review, the reviewer compared the price of the vacuum to another brand by stating "Or you can spend six or seven or eight times the money on a Dyson," and the maneuverability of the vacuum to another brand by stating, "The vacuum is a bit harder to push than our old Hoover." Thus, OTHERBRANDS is coded "2." Another vacuum cleaner reviewer stated, "It has a \$15 filter inside which sucks dirt into the plastic cannister [sic]." Because the reviewer mentioned both the filter and canister without evaluating these features, FEATURES is coded "2." A reviewer of a digital camcorder stated, "I'd hate to spend all day recording a wedding video and then lose my footage due to 'finalization failure,'" thus USAGE is coded "1" because the reviewer conveyed how he/she used the camera. A vacuum reviewer agreed with other reviewers by stating, "Like other reviewers, I was absolutely amazed by the amount of dirt I picked up on the first use," thus OTHERREVIEW1 is coded "1." Another vacuum reviewer disagreed with other reviewers by stating, "The problems that some people have mentioned in their reviews are just because they can't follow instructions, not because the machine is badly designed," thus OTHERREVIEW2 is coded "1."

3.3.2. Hypothesis testing: logistic regression analyses

We used multiple logistic regression to test H1–H6. As described below, logistic regression enables determining the change in the odds of a review being judged as helpful for a unit change in each independent variable. The dependent variable was consumers' responses to the question of whether they found the review helpful (HELP). The independent variables were the product category variable (PROD), the control variables, the variables featured in H1–H6, and the interaction between PROD and each of the control and featured variables.

Table 3 provides the logistic regression results. Significant interactions between the focal independent variables and the product type variable lend strong support to the proposed framework. Separate logistic regressions for each product category help explicate the interaction effects. Negative (positive) coefficient estimates, b, indicate that increasing the independent variable by one unit leads to lower (higher) log odds of the review being judged as helpful. The odds ratio, $\exp(b)$, is interpreted as the fractional change in the odds of a review being

[5] Remaining response categories focused on review readability. For example, 12.5% of respondents mentioned the format of the review (e.g., "I like how the review was divided into sections."), and 7.5% of respondents mentioned the use of jargon (e.g., "I couldn't understand the review because it used technical language."). The shorthand and informal linguistic culture of participants in social media and online postings renders assessments of readability non-trivial. Accordingly, we do not attempt to quantify the readability of reviews in the present data set. However, future research may wish to consider this issue more closely.

Table 1
Main study variables.

Variable type	Variable name	Description	Construct	Possible values
Dependent variable	HELP	Was the review rated as helpful?		0 = no, 1 = yes
Independent variables	PROD (H1–H6)	Product category type	Search/experience	0 = search, 1 = experience
	BALANCE (H1)	Does the review contain positive and negative evaluations of product features?	Trustworthiness	0 = no, 1 = yes
	EXPERT1 (H2(a))	Did the reviewer cite his/her direct experience with the product or cite credentials of others providing information in the review?	Expertise	0 = no, 1 = yes
	EXPERT2 (H2(b))	Did the reviewer claim "expert" status based on credentials but not experience?	Expertise	0 = no, 1 = yes
	OTHERBRANDS (H3)	Number of other brands the reviewer compared the product to	Expertise	0–no upper bound
	OTHERREVIEWS1 (H4(a))	Did the reviewer create consensus by agreeing with or clarifying other reviews of the same product?	Equivocality	0 = no, 1 = yes
	OTHERREVIEWS2 (H4(b))	Did the reviewer create confusion by disagreeing with other reviews of the same product?	Equivocality	0 = no, 1 = yes
	USAGE (H5)	Did the reviewer provide information about how he/she used the product?	Equivocality	0 = no, 1 = yes
	FEATURES (H6)	How many features does the reviewer list without evaluating?	Uncertainty	0–no upper bound
	RELATIVE[a]	INDIVIDUALEVAL–OVERALLEVAL		0 if negative 1 if positive
	POSFEATURES[a]	Number of positive evaluations of product features		0–no upper bound
	NEGFEATURES[a]	Number of negative evaluations of product features		0–no upper bound
Control variables	PRICE	Price of the reviewed product		0–no upper bound
	LENGTH	Number of words in the review		1–no upper bound
	INDIVIDUALEVAL	Numerical product evaluation by reviewer (incorporated in EVALDIFF variable)		1–5 (higher is more positive)
	OVERALLEVAL	Mean numerical product evaluation across all reviewers (incorporated in EVALDIFF variable)		1–5 (higher is more positive)
	EVALDIFF	|INDIVIDUALEVAL–OVERALLEVAL|		0–4

[a] Included in the main study supplemental analyses.

perceived as helpful when the independent variable increases by one unit (or when a dummy-coded variable changes from "0" to "1"). A ratio greater than one indicates that the factor enhances helpfulness, while an odds ratio less than one indicates that it reduces helpfulness. For example, when $b = .30$, $\exp(b) = 1.35$, indicating a 35% increase in the odds of a review being rated as helpful (versus not helpful) when the independent variable associated with b increases by one unit. Alternatively, when $b = -.30$, $\exp(b) = .74$, indicating a 26% decrease in the odds of a review being rated as helpful (versus not helpful) when the independent variable associated with b increases by one unit. A coefficient estimate close to 0 and, consequently, $\exp(b)$ close to 1 indicates that the independent variable has little effect on helpfulness.

Table 2
Main study summary statistics.

Variable	Overall sample	Experience products	Search products
HELP[a]	85.65 (19.69)	82.03 (25.20)	89.27 (10.87)
PRICE ($)[b]	464.74 (635.60)	146.09 (148.70)	783.41 (764.88)
LENGTH[b]	289.59 (295.92)	220.49 (211.52)	358.69 (348.89)
INDIVIDUALEVAL[b]	3.86 (1.48)	3.56 (1.66)	4.17 (1.20)
OVERALLEVAL[b]	4.05 (.65)	4.13 (.68)	3.96 (.61)
EVALDIFF[b]	1.03 (.91)	1.22 (1.10)	.83 (.60)
BALANCE[c]	34.44%	38.89%	30.00%
EXPERT1[c]	23.33%	28.89%	17.78%
EXPERT2[c]	3.33	1.11	5.56
OTHERBRANDS[b]	.18 (.29)	.16 (.28)	.21 (.30)
OTHERREVIEWS1[c]	8.33%	8.89%	7.78%
OTHERREVIEWS2[c]	6.11%	5.56%	6.67%
USAGE[c]	66.11%	62.22%	70.00%
FEATURES[b]	.15 (.24)	.17 (.26)	.14 (.22)
POSFEATURES[b]	.33 (.27)	.30 (.27)	.36 (.26)
NEGFEATURES[b]	.17 (.24)	.17 (.25)	.17 (.24)

[a] Mean percentage of reviews evaluated as "helpful." Standard deviation in parentheses.
[b] Mean values. Standard deviation in parentheses.
[c] Percentage of reviews exhibiting characteristic.

As shown in Table 3, the BALANCE × PROD interaction is significantly positive, revealing a stronger effect of BALANCE for experience goods than for search goods and supporting H1. Balanced (versus one-sided) reviews enhance the odds of a review being perceived as helpful by a factor of 1.54 (54%) for experience goods, but only by an insignificant factor of 1.04 (4%) for search goods.

Both the EXPERT1 and EXPERT1 × PROD effects are significant, with the significant positive interaction revealing a stronger effect for experience goods than for search goods and supporting H2(a). Claims of expertise based on experience or on the credentials of others significantly increase the odds of a review being perceived as helpful by a factor of 1.58 (58%) for search goods and by a factor of 2.81 (181%) for experience goods.

In support of H2(b), EXPERT2 has a significant negative effect for both search and experience goods. The effect is stronger for experience goods than for search goods as revealed by the significant EXPERT2 × PROD interaction. Claiming expertise based on one's own credentials reduces the odds of a review being perceived as helpful by a factor of .39 (61%) for search goods and by a factor of .03 (97%) for experience goods.

The OTHERBRANDS × PROD interaction is significantly positive, revealing that referencing other brands has a stronger effect for experience goods than for search goods and supporting H3.[6] Referring to other brands significantly increases the odds of a review being perceived as helpful by a factor of 13.20 (1220%) for experience goods but decreases the odds of a review being perceived as helpful by an insignificant factor of .99 (1%) for search goods.

The OTHERREVIEWS1 × PROD interaction is significantly positive, revealing a stronger effect for experience goods than for search goods and supporting H4(a). Referring to other reviews in ways that build consensus increases the odds of a review being perceived as helpful

[6] For comparability with the dummy variables, the continuous variables OTHERBRANDS, FEATURES, NEGFEATURES, POSFEATURES, EVALDIFF, LENGTH, and PRICE were normalized to a 0–1 scale as follows: (variable value − minimum value)/(maximum value − minimum value).

Table 3
Main study: multiple logistic regression results.

	Hypothesis	Log odds of helpful review			Odds of helpful review		
		Full model	Search	Experience	Full model	Search	Experience
Independent variable effects							
INTERCEPT		3.07^{***}	3.07^{***}	1.84^{***}	21.59	21.59	6.30
PROD		-1.23^{***}	-1.35^{***}	-1.09^{***}	.29	.26	.34
BALANCE		$.04^{NS}$	$.04^{NS}$	$.43^{***}$	1.04	1.04	1.54
EXPERT1		$.46^{**}$	$.46^{**}$	1.03^{***}	1.58	1.58	2.81
EXPERT2		$-.95^{***}$	$-.96^{***}$	-3.59^{***}	.39	.39	.03
OTHERBRANDS		$-.01^{NS}$	$-.01^{NS}$	2.58^{***}	.99	.99	13.20
OTHERREVIEWS1		$.50^{*}$	$.50^{**}$	1.78^{***}	1.65	1.65	5.93
OTHERREVIEWS2		$-.59^{*}$	$-.59^{**}$	-1.39^{***}	.55	.55	.25
USAGE		$.47^{***}$	$.47^{***}$	1.00^{**}	1.60	1.60	2.72
FEATURES		2.21^{***}	2.21^{***}	$-.48^{*}$	9.10	9.10	.62
BALANCE × PROD	H1: supported	$.39^{*}$			1.48		
EXPERT1 × PROD	H2(a): supported	$.58^{*}$			1.78		
EXPERT2 × PROD	H2(b): supported	-2.63^{***}			.07		
OTHERBRANDS × PROD	H3: supported	2.59^{***}			13.39		
OTHERREVIEWS1 × PROD	H4(a): supported	1.29^{***}			3.63		
OTHERREVIEWS2 × PROD	H4(b): supported	$-.80^{*}$			.45		
USAGE × PROD	H5: supported	$.53^{**}$			1.71		
FEATURES × PROD	H6: supported	-2.69^{***}			.07		
Control variable effects							
PRICE		-1.35^{***}	-1.35^{***}	-1.09^{***}	.26	.26	.34
LENGTH		$-.37^{NS}$	$-.37^{NS}$	$.44^{NS}$	.69	.69	1.56
EVALDIFF		-1.22^{***}	-1.22^{***}	-2.20^{***}	.30	.28	.11
PRICE × PROD		$.27^{NS}$			1.31		
LENGTH × ROD		$.82^{NS}$			2.27		
EVALDIFF × PROD		$-.99^{***}$			.37		

Notes: The full model statistically tests the hypotheses. Results for each product type aid in interpretation. The log odds columns contain model coefficients. The odds columns are interpreted as the change in the odds of a review being rated as helpful when the independent variable changes by one unit. Values < 1.0 indicate decreasing odds and values > 1.0 indicate increasing odds.

*** p < .01.
** p < .05.
* p < .10.
^{NS} p ≥ .10.

by a factor of 1.65 (65%) for search goods and by a factor of 5.93 (493%) for experience goods.

The OTHERREVIEWS2 main effect and OTHERREVIEWS2 × PROD interaction are significantly negative, supporting H4(b). Referring to other reviews in disagreeable ways significantly reduces the odds of a review being perceived as helpful by a factor of .55 (45%) for search goods and by a factor of .25 (75%) for experience goods.

Both the USAGE main effect and the USAGE × PROD interaction are significantly positive, supporting H5. Descriptions of usage situations significantly increase the odds of a review being helpful by a factor of 1.60 (60%) for search goods and by a factor of 2.72 (172%) for experience goods.

The FEATURES main effect is significantly positive, and the FEATURES × PROD interaction is significantly negative. In support of H6, listing product features significantly increases the odds of a review being perceived as helpful by a factor of 9.10 (810%) for search goods and reduces the odds of a review being perceived as helpful by an insignificant factor of .62 (38%) for experience goods.

As shown in Table 3, there is a significant, negative interaction between the control variable EVALDIFF and PROD. Specifically, the odds of an individual review being perceived as helpful when its rating deviates by one unit from the mean rating diminishes by a factor of .11 (89%) for experience goods but only by a factor of .28 (72%) for search goods. Thus, though not explicitly hypothesized, these results replicate the findings of Mudambi and Schuff [61].

3.3.3. *Supplemental analyses: unique effects of positive and negative information*

As a final consideration, we examine the effects of information valence. While Pretest 2 respondents mentioned this factor as potentially contributing to helpfulness, valence effects on helpfulness are not clear. Although some researchers claim support for a negativity bias (i.e., shoppers prefer negative reviews to positive reviews), other researchers have obtained results that call into question the robustness of a negativity bias (e.g., [71,87,88,91]). To contribute to this discussion, we assess the unique effects of negative and positive information by extracting valence from the BALANCE and EVALDIFF variables used in the previously presented logistic regression analysis. First, we created the dummy variable RELATIVE that took the value 0 when the product star rating in the review (INDIVIDUALEVAL) was lower than the average (OVERALLEVAL) product rating (i.e., a negative review) and 1 when the reviewer gave the product more stars than average (i.e., a positive review).[7] We included this variable in a logistic regression along with the product category variable (PROD) and the interaction between these variables. RELATIVE had a significant positive effect on perceived helpfulness for search goods ($b = 1.38$, Wald $\chi^2(1) = 111.07$, $p < .01$, $\exp(b) = 3.98$); when the product received more stars than average, the review was viewed as more helpful than when the product received fewer stars than average. Further, the RELATIVE × PROD interaction was positive and significant ($b = .88$, Wald $\chi^2(1) = 20.16$, $p < .01$, $\exp(b) = 2.40$), indicating that the effect was stronger for experience goods than for search goods.[8] Thus, shoppers displayed a positivity bias by indicating a general preference for positive reviews over negative reviews.

7 We excluded cases in which the review received the same number of stars as the average review; that is, INDIVIDUALEVAL = OVERALLEVAL.

8 Similar results were obtained when we ran this analysis with positive reviews (i.e., products receiving 4 or 5 stars) and negative reviews (i.e., products receiving 1 or 2 stars).

Second, we included the NEGFEATURES and POSFEATURES variables in a logistic regression, along with the interactions between these variables and the product category variable. Positive evaluative statements enhanced helpfulness for search goods ($b = 2.01$, Wald $\chi^2(1) = 62.13$, $p < .01$, $\exp(b) = 7.49$), and the significant POSFEATURES × PROD interaction indicated a stronger positive effect for experience goods ($b = 3.41$, Wald $\chi^2(1) = 72.02$, $p < .01$, $\exp(b) = 30.16$). Negative evaluative statements did not affect helpfulness for search goods ($b = -.05$, Wald $\chi^2(1) = .06$, $p = .80$, $\exp(b) = .95$), and the significant NEGFEATURES × PROD interaction indicated a *negative* effect on helpfulness for experience goods ($b = -.56$, Wald $\chi^2(1) = 4.20$, $p = .04$, $\exp(b) = .57$). These results are consistent with those based on star ratings (i.e., a positivity rather than negativity effect).

4. Discussion

4.1. Summary of results

Pretest 1 empirically verified our selection and categorization of search and experience products. Pretest 2 guided the selection of review characteristics on which the research hypotheses were based, thus providing the first direct *consumer* insight into the factors that drive helpfulness (as opposed to factors that *researchers* think are important to consumers). The main study tested the hypotheses, and Table 3 summarizes the results. When considering the results, very different pictures of helpful product reviews emerge for search and experience goods. Balanced reviews with negative and positive comments, citing one's own experience with the product or the credentials of others who provided information contained in the review, comparing the product to other brands, creating consensus with other reviews, and describing how the product had been used all enhanced helpfulness more for experience goods than for search goods. Listing, but not evaluating, features enhanced helpfulness more for search goods than for experience goods. Citing one's credentials without citing one's experience with the product and disagreeing with other reviews both adversely affected helpfulness, and these effects were stronger for experience goods than for search goods. Further, consumers were more likely to perceive as helpful reviews that (1) gave products higher than average star ratings (compared to reviews that gave lower than average star ratings) and (2) contained more positive evaluative statements (compared to reviews that contained fewer positive statements), and these effects were stronger for experience than for search goods. Interestingly, negative evaluative statements did not affect perceived helpfulness for search goods and had a negative effect on helpfulness for experience goods.

4.2. Contributions to the literature

Our research adds to the emerging body of IS and Marketing literature regarding consumer generated content. By developing a framework that combines search/experience product characteristics with aspects arising from word-of-mouth (i.e., source expertise and trustworthiness) and consumer information requirements (i.e., uncertainty and equivocality reduction), our research leads to a richer understanding of the review factors that consumers perceive as helpful as they go through their decision-making process. By refining our frameworks through such research, we can facilitate better analysis of consumer content and its efficacy, eventually leading to better decision support tools for consumers and better analytics for companies. In this section, we discuss how our findings support, challenge, and extend existing research.

With Pretest 1, we develop a rigorous product classification technique. Research that utilizes product categorization frameworks tends to base classification either on intuitive assumptions about which category a given product belongs to or on direct measures. We augment intuition and direct measures with multidimensional scaling, thereby enhancing the validity of the resulting product categories. Pretest 2 reveals which review characteristics consumers consider relevant in judging helpfulness. Researchers have assumed that certain characteristics are important, but, to our knowledge, no research has obtained direct insight into which characteristics consumers actually focus on.

The main study supports previous findings that it is essential to account for product type when examining the effectiveness of online product reviews (e.g., [36,61,71,74,87]). Importantly, while our findings corroborate those of Mudambi and Schuff [61], we provide an expanded perspective on the factors that contribute to helpfulness. For example, while Mudambi and Schuff [61] consider effects of review length, we extend their findings by showing that characteristics that contribute to length, including the listing of product features, describing usage situations, referring to other brands and reviews, reviewer claims of expertise, and listing both positive and negative product aspects, affect helpfulness. Though not hypothesized, our findings replicate the review extremity findings of Mudambi and Schuff [61], as extreme reviews reduce helpfulness, and this effect is stronger for experience goods than for search goods. We also extend their findings by examining positive versus negative reviews. The data revealed a positivity bias for both search and experience goods. Thus, our findings contribute to the growing literature that reports inconsistent and complex effects of negative and positive review information, effects that may depend on the body of reviews that a shopper reads (e.g., [55,71,87,88,91]). Given these findings, we encourage research into when and why online consumers give greater consideration to positive or negative information. Such research may benefit from incorporating the notion of cue congruence, in which consumer preference for information depends on the congruence between consumption goals and consumption-related cues [36]. Research may also benefit from considering differential attributions that consumers make about negative and positive information (e.g., [50]), and from more general findings that people broadly overestimate their effectiveness as communicators due to an (egocentric) inability to fully ignore their own phenomenology and internal goals when constructing messages for others, who are unlikely to share the same phenomenology and goals (e.g., [44,45,48,67]).

While trust is a critical issue in online shopping (e.g., [2,3,46,47,80]), how online shoppers judge the expertise of reviewers is an area in need of more research. For example, the mere provision of a peer recommendation can affect consumer choice regardless of the reviewer's profile or credibility [74]. However, identity-descriptive information can enhance the rating of product reviews [25]. Expanding on this insight, we find that *how* one claims expertise affects judgments of helpfulness. Reviewers can enhance helpfulness by describing their direct experience with the product. However, when a reviewer simply cites his/her credentials, but does not provide evidence of first-hand experience, consumers find the review less helpful than if the reviewer does not provide credentials. One explanation for these findings comes from research showing that displays of excessive pride, or hubris, adversely affect favorably ratings [37,79]. Thus, consumers may punish a reviewer by evaluating his/her review as unhelpful if the reviewer seems to take too much pride in his/her accomplishments. We also note the possibility that expertise claims may reduce the ability of the reader to relate to the reviewer. For example, a reviewer of a DVD player stated, “Based on my many years of audio and video usage on the higher end of the market …” If readers of this review do not consider themselves “high end” audio/video users, they may not find the review helpful. While our data do not allow us to test the effects of reader-reviewer fit, we encourage researchers to consider this possibility in efforts to more clearly explicate the conditions under which information about the reviewer enhances perceptions of helpfulness.

Further, we find that helpfulness is differentially impacted by how one references other reviews. In support of our predictions, reviewers enhance helpfulness by agreeing with another review or explaining conflicting reviews. However, reviewers adversely affect helpfulness when they disagree with another review. Interestingly, although consumers who post reviews often do so with the goal of appearing thoughtful and helpful [18], reviewers may be blind to the possibility that certain ways of stating expertise or referring to other reviews can have adverse effects on helpfulness, or that certain types of information can have minimal or no effects on helpfulness (e.g., referencing other brands for search goods, listing product features for experience goods). Such unintentional unhelpfulness may arise due to reviewers overweighting their own preferences when trying to predict the information preferences of others (see [67], p. 48 for a summary of studies that suggest this possibility).

4.3. Pragmatic implications for systems

The results of this research can assist companies in developing better decision support systems for consumers, as well as engaging in useful analytics that can enhance customer value. Currently, the low cost of providing information online leads retailers to provide consumers with access to *all* reviews. However, too much information can lead to poorer decisions by overwhelming shoppers [40]. Today's aggregated review information typically involves simple averages of consumer ratings and/or rudimentary search capabilities. Thus, to get a good qualitative assessment of the product, customers must often wade through pages of textual information. However, by knowing the type of information consumers' desire, retailers can potentially lower costs by being more efficient communicators. In particular, retailers may be able to use consumer-created product information to complement or replace retailer-created information [11]. Further, provision of decision support that allows consumers to quickly sort through reviews based on more sophisticated criteria, such as whether the reviewer has hands-on experience with the product, lists product features, provides balanced reviews, or compares with other brands, would be extremely useful to consumers by allowing them to narrow down reviews to the ones they might consider most helpful. The current research suggests that a simple input from the user regarding their perception of the search or experience nature of the product can usefully refine these filtering mechanisms even further. Systems that can conduct sophisticated semantic textual analysis are well within the purview of contemporary technologies. With a well-designed interface, such knowledge can also help retailers develop their websites to be more user-friendly with improved design of online review systems to catch consumer attention. Further enhancement of these concepts could also lead to more sophisticated systems where specific product recommendations made by the system are buttressed with selective data pulled from “helpful” reviews.

Greater understanding of the drivers of perceived helpfulness may also aid in conducting analytics to obtain superior information in planning processes. For example, in alignment with our framework for understanding perceptions of review helpfulness, marketers can effectively segment and profile customers according to perceptions of the benefits and uncertainty of shopping online [7]. Similarly, incorporating information found in online reviews can improve sales forecasts [19]. Recent work in IS may help accelerate the integration of reviews into marketing planning, as artificial intelligence and text mining tools allow for automation of content analysis and reviewer rankings (e.g., [29,90]). Tools can also automatically fill in missing helpfulness ratings based on such analysis. Our findings can inform the design, training, and calibration of such tools, which rely heavily on assumptions about consumers' use and perception of product-related language and information.

Firms are often able to manage, to some degree, social interactions between consumers [32]. Along these lines, our findings suggest that retailers may want to consider providing standardized review forms for shoppers to complete, potentially serving to make reviews more helpful to shoppers. Such forms may have two benefits. First, they could increase the chance that reviewers provide critical and relevant information, as the forms would serve to guide reviewers. That is, retailers could create forms that request specific information. Some retailers are already doing this. For example, Lowes.com asks reviewers to indicate their level of expertise (though our findings suggest the potential for this to backfire). Amazon.com provides tips for writing good reviews, including encouraging reviewers to share their experiences with, or usage of, the product (consonant with our finding that usage information is valuable for search and experience goods). Importantly, review forms could differ by product, depending on the information that prior shoppers found most helpful. Second, these forms may improve the layout and navigability of the resulting reviews, potentially enhancing ease of use and customer satisfaction [33,84,89]. In support of this latter benefit, in Pretest 2, a number of respondents (12.5% on average across product categories) mentioned that review format impacted helpfulness. Relatedly, companies sometimes wish to influence shoppers' product opinions in online forums through “promotional chat” and other means [17,56]. Today's social media forums could facilitate this kind of engagement. Our findings provide guidance for content managers who wish to contribute product review information that is perceived as particularly helpful by consumers for particular types of products.

Finally, our findings that reviewers sometimes provide information that has no, or even detrimental, effects on helpfulness, and the possibility that egocentrism makes reviewers unaware of their unhelpfulness as communicators [44,45,48,67], suggests that better review forms alone may be insufficient for improving product reviews. Instead, websites eliciting reviews may need to change the mindsets of review writers to reduce reliance on their own knowledge and information preferences. Although challenging, Pronin and Kugler [68] find that educating people about introspection bias can overcome the effect and its resulting problems. We encourage research into techniques for achieving this goal in the context of writing online reviews.

4.4. Conclusion

Our work enriches understanding of online reviews and draws implications for how companies can better leverage these reviews and design superior online decision support systems. It extends prior work in this area both theoretically and methodologically. We acknowledge that there may be alternative and equally productive methods for analyzing the content of reviews. For example, researchers with expertise in linguistics or discourse theory may examine reviews from semantic, pragmatic, spatial, and temporal perspectives to learn more about when and how consumers use different types of language in reviews. Also, while consumers generally judged the reviews in our sample to be helpful, relatively unhelpful reviews may also impact consumers' attitudes and responses to products, channels, or retailers. In all cases, given the importance of catching consumer attention in this increasingly information laden world, it is critical to continue to assess and evaluate information content of reviews in order to derive implications for designing websites that enhance outcomes for customers, manufacturers, retailers, and other stakeholders.

References

[1] E. Anderson, D. Simester, Reviews without a purchase: low ratings, loyal customers and deception, J. Mark. Res. 51 (3) (2014) 249–269.
[2] N.F. Awad, A. Ragowsky, Establishing trust in electronic commerce through online word of mouth: an examination across genders, J. Manag. Inf. Syst. 24 (4) (Spring 2008) 101–121.
[3] S. Ba, P.A. Pavlou, Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior, MIS Q. 26 (3) (September 2002) 243–268.
[4] H. Baek, J. Ahn, Y. Choi, Helpfulness of online consumer reviews: readers' objectives and reviewer cues, Int. J. Electron. Commer. 17 (2) (Winter 2012–2013) 99–126.
[5] R. Batra, I. Sinha, Consumer-level factors moderating the success of private label brands, J. Retail. 76 (2) (Summer 2000) 175–191.

[6] W.O. Bearden, R.G. Netemeyer, J.E. Teel, Measurement of consumer susceptibility to interpersonal influence, J. Consum. Res. 15 (March 1989) 473–481.
[7] A. Bhatnagar, S. Ghose, Segmenting consumers based on the benefits and risks of internet shopping, J. Bus. Res. 57 (12) (December 2004) 1352–1360.
[8] P.F. Bone, Word-of-mouth effects on short-term and long-term product judgments, J. Bus. Res. 32 (3) (March 1995) 213–223.
[9] I. Borg, P.J.F. Groenen, Modern Multidimensional Scaling: Theory and Applications, Springer, New York, 1997.
[10] M. Cave, Market models and consumer protection, J. Consum. Policy 8 (4) (December 1985) 334–350.
[11] Y. Chen, J. Xie, Online consumer review: word-of-mouth as a new element of marketing communication mix, Manag. Sci. 54 (3) (March 2008) 477–491.
[12] J. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, J. Mark. Res. 43 (3) (August 2006) 345–354.
[13] V. Choudhury, E. Karahanna, The relative advantage of electronic channels: a multi-dimensional view, MIS Q. 31 (1) (March 2008) 179–200.
[14] E.K. Clemons, G.G. Guodong, L.M. Hitt, When online reviews meet hyperdifferentiation: a study of the craft beer industry, J. Manag. Inf. Syst. 23 (2) (2006) 149–171.
[15] R.L. Daft, R.H. Lengel, Organizational information requirements, media richness, and structural design, Manag. Sci. 32 (5) (May 1986) 554–571.
[16] R.L. Daft, N.B. Macintosh, A tentative exploration into the amount and equivocality of information processing in organizational work units, Adm. Sci. Q. 26 (1981) 207–224.
[17] C. Dellarocas, Strategic manipulation of internet opinion forums: implications for consumers and firms, Manag. Sci. 52 (10) (October 2006) 1577–1593.
[18] C. Dellarocas, G.G. Gao, R. Narayan, Are consumers more likely to contribute online reviews for hit or niche products, J. Manag. Inf. Syst. 27 (2) (Fall 2010) 127–158.
[19] C. Dellarocas, X. Zhang, N.F. Awad, Exploring the value of online product reviews in forecasting sales: the case of motion pictures, J. Interact. Mark. 21 (4) (Autumn 2007) 23–45.
[20] J. De Leeuw, W.J. Heiser, Multidimensional scaling with restrictions on the configuration, in: P. Krishnaiah (Ed.)Multivariate Analysis, vol. V, North Holland Publishing Company, Amsterdam, The Netherlands 1980, pp. 501–522.
[21] W. Duan, B. Gu, A.B. Whinston, Do online reviews matter?: An empirical investigation of panel data. *Decision Support Systems*, 45 (4) (2008) 1007–1016.
[22] T. Erdem, J. Swait, Brand equity as a signaling phenomenon, J. Consum. Psychol. 7 (2) (1998) 131–157.
[23] K.A. Ericsson, M.J. Prietula, E.T. Cokely, The making of an expert, Harv. Bus. Rev. (July) (2007).
[24] K. Floyd, R. Freling, S. Alhoqail, H.Y. Cho, T. Freling, How online product reviews affect retail sales: a meta-analysis, J. Retail. 90 (2) (2014) 217–232.
[25] C. Forman, A. Ghose, B. Wiesenfeld, Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, Inf. Syst. Res. 19 (3) (September 2008) 291–313.
[26] D. Gefen, E. Karahanna, D.W. Straub, Trust and the TAM in online shopping: an integrated model, MIS Q. 27 (1) (March 2003) 51–90.
[27] D.H. Gensch, A two-stage disaggregate attribute choice model, Mark. Sci. 6 (3) (Summer 1987) 223–231.
[28] A.D. Gershoff, A. Mukherjee, A. Mukhopadhyay, Consumer acceptance of online agent advice: extremity and positivity effects, J. Consum. Psychol. 13 (1/2) (2003) 161–170.
[29] A. Ghose, P.G. Ipeirotis, Towards an understanding of the impact of customer sentiment on product sales and review quality, Proceedings of the 2006 Workshop on Information Technology and Systems, 2006.
[30] F. Glemet, R. Mira, The brand leader's dilemma, McKinsey Q. 2 (1993) 3–15.
[31] D. Godes, D. Mayzlin, Using online conversations to study word of mouth communication, Mark. Sci. 23 (4) (Fall 2004) 545–560.
[32] D. Godes, D. Mayzlin, Y. Chen, S. Das, C. Dellarocas, B. Pfeiffer, B. Libai, S. Sen, M. Shi, P. Verlegh, The firm's management of social interactions, Mark. Lett. 16 (3/4) (December 2005) 415–428.
[33] R. Grant, R.J. Clarke, E. Kyriazis, A review of factors affecting online consumer search behaviour from an information value perspective, J. Mark. Manag. 23 (5/6) (July 2007) 519–533.
[34] G. Häubl, V. Trifts, Information availability and consumer preference: can online retailers benefit from providing access to competitor price information, J. Consum. Psychol. 13 (1/2) (2003) 149–159.
[35] T. Hennig-Thurau, K.P. Gwinner, G. Walsh, D.D. Gremler, Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? J. Interact. Mark. 18 (1) (Winter 2004) 38–52.
[36] D.L. Hoffman, T.P. Novak, F. Wan, The impact of online product review characteristics on consumer preferences, Adv. Consum. Res. 31 (2003) 530–531.
[37] V. Hoorens, M. Pandelaere, F. Oldersma, C. Sedikides, The Hubris hypothesis: you can self-enhance, but you'd better not show it, J. Pers. 80 (5) (2012) 1237–1274.
[38] C.I. Hovland, I.K. Janis, H.H. Kelley, Communication and Persuasion, Yale University Press, New Haven, C.T., 1953
[39] N. Hu, J. Zhang, P. Pavlou, Overcoming the J-shaped distribution of product reviews, Commun. ACM 52 (10) (October 2009).
[40] J. Jahng, H. Jain, K. Ramamurthy, Effective design of electronic commerce environments: a proposed theory of congruence and an illustration, IEEE Syst. Man Cybern. 30 (4) (July 2000) 456–471.
[41] Z. Jiang, I. Benbasat, The effects of presentation formats and task complexity on online consumers' product understanding, MIS Q. 31 (3) (September 2007) 475–500.
[42] F.R. Jimenez, N.A. Mendoza, Too popular to ignore: the influence of online reviews on purchase intentions of search and experience products, J. Interact. Mark. 27 (2013) 226–235.
[43] M.A. Kamins, H. Assael, Two-sided versus one-sided appeals: a cognitive perspective on argumentation, source derogation, and the effect of disconfirming trial on belief change, J. Mark. Res. 24 (1) (February 1987) 29–39.
[44] C.L. Kawada, G. Oettingen, P.M. Gollwitzer, J.A. Bargh, The projection of implicit and explicit goals, J. Pers. Soc. Psychol. 86 (4) (2004) 545–559.
[45] B. Keysar, A.S. Henly, Speakers' overestimation of their effectiveness, Psychol. Sci. 13 (3) (2002) 207–212.
[46] D. Kim, D.L. Ferrin, R. Rao, Trust and satisfaction, two stepping stones for successful E-commerce relationships: a longitudinal exploration, Inf. Syst. Res. 20 (2) (June 2009) 237–257.
[47] D. Kim, I. Benbasat, Trust-assuring arguments in B2C E-commerce: impact of content, price, and source on trust, J. Manag. Inf. Syst. 26 (3) (Winter 2009-10) 175–206.
[48] J. Kruger, N. Epley, J. Parker, Z.W. Ng, Egocentrism over e-mail: can we communicate as well as we think, J. Pers. Soc. Psychol. 89 (6) (2005) 925–936.
[49] N. Kumar, I. Benbasat, The influence of recommendations and consumer reviews on evaluations of websites, Inf. Syst. Res. 17 (4) (December 2006) 425–439.
[50] R.N. Laczniak, T.E. DeCarlo, S.N. Ramaswami, Consumers' responses to negative word-of-mouth communication: an attribution theory perspective, J. Consum. Psychol. 11 (1) (2001) 57–73.
[51] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (March 1977) 159–174.
[52] M. Li, L. Huang, C.H. Tan, K.K. Wei, Helpfulness of online product reviews as seen by consumers: source and content features, Int. J. Electron. Commer. 17 (4) (Summer 2013) 101–136.
[53] K.H. Lim, I. Benbasat, The effect of multimedia on perceived equivocality and perceived usefulness of information systems, MIS Q. 24 (3) (September 2000) 449–471.
[54] Y. Liu, Word-of-mouth for movies: its dynamics and impact on box office revenue, J. Mark. 70 (3) (July 2006) 74–89.
[55] S. Ludwig, K. de Ruyter, M. Friedman, E.C. Bruggen, M. Wetzels, G. Pfann, More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates, J. Mark. 77 (January 2013) 87–103.
[56] D. Mayzlin, Promotional chat on the internet, Mark. Sci. 25 (2) (Spring 2006) 155–163.
[57] D. Mayzlin, Y. Dover, J. Chevalier, Promotional reviews: an empirical investigation of online review manipulation, Am. Econ. Rev. 104 (8) (2014) 421–2455.
[58] D.H. McKnight, V. Choudhury, C. Kacmar, Developing and validating trust measures for e-commerce: an integrative typology, Inf. Syst. Res. 13 (3) (September 2002) 334–359.
[59] B. Mittal, Lack of attribute searchability: some thoughts, *Psychology & Marketing*, 21 (6) (June 2004) 443–462.
[60] W.W. Moe, M. Trusov, The value of social dynamics in online product ratings forums, J. Mark. Res. 48 (3) (2011) 444–456.
[61] S.M. Mudambi, D. Schuff, What makes a helpful online product review? A study of customer reviews on Amazon.com, MIS Q. 34 (1) (March 2010) 185–200.
[62] P. Nelson, Advertising as information, J. Polit. Econ. 82 (4) (July/August 1974) 729–754.
[63] R. Ohanian, Construction and validation of a scale to measure celebrity endorsers' perceive expertise, trustworthiness, and attractiveness, J. Advert. 19 (3) (1990) 39–52.
[64] Y. Pan, J.Q. Zhang, Born unequal: a study of the helpfulness of user-generated product reviews, J. Retail. 87 (4) (2011) 598–612.
[65] C. Pechmann, Predicting when two-sided ads will be more effective than one sided ads: the role of correlation and correspondent inferences, J. Mark. Res. 29 (4) (1992) 441–453.
[66] D.M. Phillips, J.C. Olson, H. Baumgartner, Consumption visions in consumer decision making, Adv. Consum. Res. 22 (1995) 280–284.
[67] E. Pronin, The introspection illusion, in: Mark P. Zanna (Ed.) Advances in Experimental Psychology, vol. 41, Academic Press, Burlington 2009, pp. 1–67.
[68] E. Pronin, M.B. Kugler, Valuing thoughts, ignoring behavior: the introspection illusion as a source of the bias blind spot, J. Exp. Soc. Psychol. 43 (2007) 565–578.
[69] N. Purnawirawan, P. De Pelsmacker, N. Dens, Balance and sequence in online reviews: how perceived usefulness affects attitudes and intentions, J. Interact. Mark. 26 (2012) 244–255.
[70] J.A. Rosa, A.J. Malter, E-(embodied) knowledge and e-commerce: how physiological factors affect on-line sales of experiential products, J. Consum. Psychol. 13 (1/2) (2003) 63–75.
[71] S. Sen, D. Lerman, Why are you telling me this? An examination into negative consumer reviews on the web, J. Interact. Mark. 21 (4) (Autumn 2007) 76–94.
[72] S. Senecal, J. Nantel, The influence of online product recommendations on consumers' online choices, J. Retail. 80 (2) (Summer 2004) 159–169.
[73] J.J. Skowronski, D.E. Carlston, Social judgment and social memory: the role of cue diagnosticity in negativity, positivity, and extremity biases, J. Pers. Soc. Psychol. 52 (4) (1987) 689–699.
[74] D. Smith, S. Menon, K. Sivakumar, Online peer and editorial recommendations, trust, and choice in virtual markets, J. Interact. Mark. 19 (3) (Summer 2005) 15–37.
[75] K.J. Srnka, S.T. Koeszegi, From words to numbers: how to transform qualitative data into meaningful quantitative results, Schmalenbach Bus. Rev. 59 (1) (January 2007) 29–57.
[76] M. Sun, How does the variance of product ratings matter, Manag. Sci. 58 (4) (2012) 696–707.
[77] K.W.P. Tan, D. Swee, C. Lim, B.H. Detenber, L. Alsagoff, The impact of language variety and expertise on perceptions of online political discussions, J. Comput.-Mediat, Commun. 13 (1) (2008) 76–99.
[78] N. Tractinsky, V.S. Rao, Incorporating social dimensions in web-store design, Hum. Syst. Manag. 20 (2) (2001) 105–121.

[79] J.L. Tracy, R.W. Robins, The psychological structure of pride: a tale of two facets, J. Pers. Soc. Psychol. 92 (3) (2007) 506–525.
[80] O. Turel, Y. Yuan, C.E. Connelly, In justice we trust: predicting user acceptance of e-customer services, J. Manag. Inf. Syst. 24 (4) (Spring 2008) 123–151.
[81] J. Voight, Getting a handle on customer reviews, Adweek 48 (26) (2007) 16–17.
[82] B.A. Walker, J.C. Olson, The activated self and consumer behavior: a cognitive structure perspective, in: Russell W. Belk (Ed.) Research in Consumer Behavior, vol. 8, JAI Press, Greenwich, C.T. 1994, pp. 135–171.
[83] D. Weathers, S. Sharma, S.L. Wood, Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods, J. Retail. 83 (4) (December 2007) 393–401.
[84] B.D. Weinberg, P.D. Berger, R.C. Hanna, A belief-updating process for minimizing waiting time in multiple waiting-time events: application in website design, J. Interact. Mark. 17 (4) (Autumn 2008) 24–37.
[85] A.M. Weiss, N.H. Lurie, D.J. Macinnis, Listening to strangers: whose responses are valuable, how valuable are they, and why, J. Mark. Res. 45 (August 2008) 425–436.
[86] B. Wernerfelt, Efficient marketing communication: helping the customer learn, J. Mark. Res. 33 (2) (May 1996) 239–246.
[87] L.M. Willemsen, P.C. Neijens, F. Bronner, J.A. de Ridder, 'Highly recommended!' The content characteristics and perceived usefulness of online consumer reviews, J. Comput.-Mediat. Commun. 17 (2011) 19–38.
[88] P.F. Wu, In search of negativity bias: an empirical study of perceived helpfulness on online reviews, Psychol. Mark. 30 (11) (2013) 971–984.
[89] B. Xiao, I. Benbasat, E-commerce product recommendation agents: use, characteristics, and impact, MIS Q. 31 (1) (March 2007) 137–209.
[90] K. Xu, S.S. Liao, J. Li, Y. Song, Mining comparative opinions from customer reviews for competitive intelligence, Decis. Support. Syst. 50 (4) (2011) 743–754.
[91] D. Yin, S.D. Bond, H. Zhang, Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews, MIS Q. 38 (2) (2014) 539–560.
[92] Y. You, G.G. Vadakkepatt, A.M. Joshi, A meta-analysis of electronic word-of-mouth elasticity, J. Mark. 79 (March) (2015) 19–39.
[93] F. Zhu, X. Zhang, Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics, J. Mark. 74 (2) (March 2010) 133–148.
[94] L. Zhu, I. Benbasat, Z. Jiang, Let's shop online together: an empirical investigation of collaborative online shopping support, Inf. Syst. Res. 21 (4) (December 2010) 872–891.
[95] Y. Hu, S.S. Sundar, Effects of online health sources on credibility and behavioral intentions, Commun. Res. 37 (February 2010) 105–132.

Danny Weathers is an Associate Professor of Marketing at Clemson University.

Scott Swain is an Assistant Professor of Marketing at Clemson University.

Varun Grover is the William S. Lee (Duke Energy) Distinguished Professor of Information Systems at Clemson University.