Amazonian Forests and Trees: Multiplicity and Objectivity in Studies of Online

Consumer-Generated Ratings and Reviews

ROBERT KOZINETS*

* Robert V. Kozinets is the Hufschmid Chair of Strategic Public Relations and Business communications at the University of Southern California, email: rkozinets@usc.edu .

ABSTRACT

Consumer-generated ratings and reviews play an important role in people's experiences of online search and shopping. This paper applauds and extends the thought-provoking response of de Langhe et al. (2016) to Simonson's (2015) assertions about the topic and suggests an agenda for future research. Follow-up research into the topic should emphasize the diversity of consumers and the multiplicity of their needs. It should recognize that reviews and ratings' are complex social conversations embedded in consumers' multifaceted communicational repertoires. It should be cautious when using terms such as objective and rational when describing consumers and consumption. Being aware of the risks to external validity of studying average ratings may lead to frameworks with greater contextual integrity, and encourage collaborative communication between scholars from different perspectives working in this field.

Consumer-generated ratings and reviews play an important role in people's experiences of online search and shopping, influencing the consumption of movies, food, travel, automobiles, home services, technology, and much else. Laying out a broad course for "the BDT project", Simonson (2015, 32) proposes that researchers now focus on the interaction between consumer judgment and decision-making and the "evolving information environment", which includes researching the use of these consumer-generated ratings and reviews. Simonson (2015) makes a range of broad assertions regarding their effects, including that they: (a) have turned experience goods into search goods (21), (b) provide "rational" information about products' "harder" dimensions such as performance ratings and product use (23), (c) make "it easier for consumers to assess the quality (or "absolute value"; see Simonson and Rosen, 2014) of products, potentially making brands less important" (22), and, (d) "often represent the most reliable and most accessible predictors of product quality and user experience" (25). It is into this theoretical environment of assertions about the rational, reliable, informational value of consumer-generated rating and reviews that de Langhe, Fernbach, and Lichtenstein (2016) position their study.

The purpose of this short paper is not merely to applaud the thorough and thought-provoking response of de Langhe et al. (2016) to Simonson's (2015) assertions about the topic. In addition, I wish to respond and extend their paper, showing how it contains a range of underlying assumptions that might be tested, and setting out a succinct agenda for further research into the topic. First, this paper emphasizes the multiplicity of consumer needs. Second, it points researchers to reviews and ratings' complex social communication environment. Third, it questions some of the assumptions of the de Langhe et al. (2015) study. Finally, it provides a framework to promote appreciation of different scholarly approaches to the topic and encourage collaborative communication between them.

Appreciating the Diversity of Consumer Needs and Online Rating and Review Uses

We must base our understanding of online ratings not on assumptions of their use but upon knowledge of their actual, real-world use by consumers. De Langhe et al. (2016) want to test Simonson's (2015) assertions about quality and "harder" dimensions, and so they choose to study Amazon ratings for "relatively vertically-differentiated" product categories which "can be reliably ranked according to objective standards (e.g., electronics, appliances, power tools)" and for which "consumers typically care a lot about attributes that are objective in nature". Drawing on the power tool example, for instance, I conducted a quick netnographic (see Kozinets 2002) scan of the DeWalt's CD970K-2 18-Volk Compact Drill/Driver Kit on Amazon, which was simply the first power tool to come up in my search. What I find is an immense wealth of information about many aspects of the power tool product experience.

The different types of questions people ask about the DeWalt drill reveal a panoply of power tool perspectives. The first question asks about the battery and its charger, the second about where it is manufactured, and the third about the case. Some people want to drill into concrete and plaster, while others want to use certain types of drill bits. Some want to know if the batteries will power a flashlight, if the product will work "in Ireland", or if it comes with other options, such as a stud sensor. There are a variety of different needs and perspectives represented. Hence, even for something as apparently "vertically differentiated" as a power tool, there are many different aspects of the product itself, its packaging, case, lighting, electrical powering, service, warranty, prices, and brand image that might be salient to different people, or even to the same person at different times or for different uses.

In unpublished, ethnographic research that I conducted for a corporate client with consumers in their homes, I studied beauty product online shopping behaviors on sites that included Amazon.com. What I found was that shoppers carefully read textual reviews and examined specific, general, and average ratings. The more sophisticated, experienced, and motivated shoppers were interested in matching the product review and rating to a person who resembled themselves on some relevant dimension or dimensions—such as ethnicity, age, skin tone, eye color, hair color, or location (for seasonality concerns such as dry skin in winter). I observe the same sort of behavior on travel websites such as TripAdvisor. For example, people travelling with children of a certain age seek ratings and reviews of a hotel or destination from other people with children of the same or similar age. It would be difficult to classify this behavior as seeking an "objective" rating. On the contrary, it seems highly subjective.

As I will later develop, this multiplicity of consumer needs complicates the notion that we can cleave "objective or technical aspects of product performance" from "more subjective aspects" (cf. de Langhe et al. 2016). My beauty product shoppers were interested particularly in the performance and reliability aspects of products for their particular type of skin, face, or hair. Consumers want information that is very personal. Yet the performance characteristics they seek must be considered objective. Shoppers want to know the answer to a personal *and* performance-oriented question: "Will this product actually work *for me*, in my context? Will it do what I want it to do?" The multiplicity of consumer needs occurs because the realities of product use vary between individuals and contexts. Indeed, this is why we have segmentation, target, and positioning in marketing. Paying close attention to consumers' distinctive needs and characteristics, divergent perspectives, different subjective realities, and multifarious personal goals will help us conduct better research.

Amazon Ratings as Elements of a Cultural and Social Communication System

Reviews and ratings offer consumers a social conversation, a communications environment that they use not only to talk about the objective and subjective characteristics of products and services, but also to socialize and communicate about themselves. Kumar and Benbasat (2006) find that providing customer reviews on websites improves customer perceptions of usefulness and social presence. Klaus (2013) studied the Amazon.com customer journey and found that for one customer "Reading customer reviews is really helpful because it gives me more information about the book or product, but it is also interesting to know about the experiences of other people using the product" (448). Klaus (2013, 448) conceptualizes this aspect as "social presence," which "constitutes attributes reflecting the customer's virtual interaction with other shoppers through comments, product reviews, and social media linkages", and which "was often cited with reference to its impact on the purchase decision process, in particular in the information search and alternatives evaluation stages" on Amazon.

Indeed, much of my early work studied and built theory about consumers' production and use of online reviews of media programs, food, clothing, and others products (e.g., Kozinets 1997, 1999, 2002). Although some of these products, like television shows and shots of espresso, might be termed "experience goods", consumers were nonetheless highly motivated to describe, assess, rate, teach, develop criteria, and demonstrate their evaluative expertise about their characteristics as well as to search out, share, and debate what were the highest quality offerings in the marketplace. The behavior of creating these reviews, sharing them, teaching one another about associated products and services, commenting, and complimenting them is both cultural—it communicates and bears meaning—as well as social—it creates affiliative connection between

people. My concept of "virtual communities of consumption" (Kozinets 1999) has at its core the idea that online communications about consumption interweave pragmatic with social information: online "consumption knowledge is developed in concert with social relations" (254). Our continuing research must be constantly attuned to the social and cultural realities of consumer-generated online ratings and reviews, as well as the various social and cultural aspects of their creation, sharing, and use.

There is, in fact, a thriving and popular genre of humorous Amazon reviewing which demonstrates the rich complexity of the online rating and review environment. One of my favorite sets of reviews concerns a dedicated 1.5 meter cable made by Denon to connect DVDs and CD players to a Denon receiver. Amazon reviewers use their rating of the expensive cable to humorously extend their imaginations as well as to poke fun at people who take technical characteristics and the rating of technology products overly seriously. The following is an example which 4,091 of 4,144 people rated as helpful:

> "[Heading:] Great cable, but too fast. [Text:] Transmission of music data at rates faster than the speed of light seemed convenient, until I realized I was hearing the music before I actually wanted to play it. Apparently Denon forgot how accustomed most of us are to unidirectional time and the general laws of physics. I tried to get used to this effect but hearing songs play before I even realized I was in the mood for them just really screwed up my preconceptions of choice and free will. I'm still having a major existential hangover." ("Frank Schulze" on Amazon.com)

Consumers have the option to contribute written text, questions, answers, comments, images, photographs, ratings of products, and ratings of review helpfulness on Amazon.com. In addition, they can use the Internet to search Facebook groups, YouTube videos, Twitter posts,

podcasts, and forum discussions, as well as newspaper and magazine articles, Pinterest pins, other retails sites and their comments, or they can be brought there by hyperlinks, reviews, or other comments. They can use apps such as Yelp, Whatsapp, Instagram, and Snapchat to access reviews on the go, query friends and contacts in real time, and see photos of products in various contexts such as in other consumers' homes. Consumption of reviews and ratings occurs in this choice-saturated, diverse communication context that Madianou and Miller (2012) call *polymedia*, "within which each individual medium is defined in relational terms in the context of all other media" (170). To be thorough and do justice to the entire and actual context in which judgment and decision making occur, future research into consumer-generated reviews and ratings must shift emphasis "from a focus on the qualities of each particular medium as a discrete context" to an understanding of it as part of an entire communication repertoire used by consumers attempting to balance constraints and goals with "the ways in which interpersonal relationships are enacted and experienced" (Madianou and Miller, 170-171).

De Langhe et al. (2016) are correct in their conclusion that the Internet is not "making consumers more rational". The Internet is a remarkably complex and varied social and technological context that consumers use not only to share products' performance ratings and experiences, but also to engage, explore, connect, inform and fulfill a wide range of social, communicative, emotional, and identity-focused needs. When people use Amazon.com's review and rating system, they use it socially. Although the software often is used as a source of peer opinion and information to inform decisions about potential purchases, it also acts as a platform for cultural connection, witty repartee, social commentary, entertainment, personal revelation, self-promotion, revenge seeking, and many other activities. Consumers' use of Amazon.com, like their use of every communication interface, is always a social and cultural experience. Thus,

it is vitally important that future consumer research recognize the complexity of the contemporary communications environment in which activities such as creating and using consumer reviews and ratings are embedded.

Questioning the Assumptions of Economics, Objectivity, and Single Ratings

The term "objective quality" is rather slippery in the consumption context. Can we truly judge the absolute quality of a product like a vacuum cleaner in some objective and general sense that stands apart from the individual consumers and their differentiated needs? For some vacuum consumers, being able to easily lift and carry the vacuum cleaner up and down three flights of stairs is of paramount importance. For others, it is the ability to handle large amounts of pet hair. Others want one that will store conveniently. Vacuum cleaners also have different design elements, colors, shapes, benefits, and sizes. Consumers have different physical sensitivities, brand preferences, and cultural tastes. Each different element of a product communicates meaning and offers value differently to different people.

Thus it seems that subjective perspectives and positions will always matter to consumers' evaluations of a given product's quality. There is no universal standard upon which to base designations of real, true, or actual quality—and concomitantly no way to assess the so-called biases that allegedly detract from it. Although it may be useful for analytic purposes to create convenient fictions such as "search goods" and "vertically-differentiated product categories" that exist in a dimension of undifferentiated consumer needs, we must be careful not to let our assumptions about how the world might be influence our ability to perceive how the world truly is. De Langhe et al. (2015) assume that some set of products can be largely functional, with their

9

utility curves readily revealed, and then discover in a meticulous series of studies that: (1) price efficiently predicts quality, (2) brands are matters of costs versus benefits, and (3) consumers want to be rational decision makers. These findings bear a remarkable resemblance to the assumptions of classical economics. However, they may not be valid as descriptions of a reality where a multiplicity of different consumer needs exist.

Further, I urge caution in posing and asking research questions regarding whether consumers use Amazon ratings –or any other reviews or ratings—"appropriately". The question contains within it the deductive whispers of a forthcoming right answer. The underlying assumption of the question is that reaching a decision based upon objective ratings—as proxied by *Consumer Reports* scores—should be the goal of consumers' rating use. This assumption ignores or denigrates the many other uses to which people put ratings and reviews on Amazon.com, such as communicating with others, expressing oneself, and jokingly claiming that a cable can warp time and space. Further, it ignores the subjective filters through which people view all information.

To be realistic, future research into the topic should examine in realistic, naturalistic contexts how consumers create, share, and use ratings and reviews. How do consumers balance different types of information with the information on Amazon.com, *Consumer Reports*, and other sources? Are friends on Facebook viewed differently than experts on *Consumer Reports*? Do helpfulness ratings on Amazon.com matter? Are they suspicious of reviews that seem overly favorable, or overly critical? What are the patterns in this performance? How do these practices vary—by expertise, age, gender, education, national orientation, online site and source? Further, research might inquire into the many nonmarket and non-rational uses of reviews, such as activism, advocacy, creative communication and artistic self-expression.

10

It may be important for future researchers to appreciate that virtually all users of online ratings and reviews read and use text comments (Pavlou and Dimoka 2006). In one of the top rated reviews of the DeWalt drill/driver kit I examined earlier, "Harv" (a pseudonym, following netnographic convention), warns people about the power tool's "smaller batteries", their non-replaceable nature, its limited speed and its workload limitations. However, Harv still gives the drill 5 stars, explaining in his review that he did so because "it finished the job with a lot of coaxing and still works despite melting something internally." Another top-rated review, by "Dave", waxes enthusiastic about the brand, tells us about the longevity of his last Dewalt drill, rates the new one "not a bad drill", then complains about the Dewalt website and online contact form as well as the product's batteries. Dave rates the drill one star out of five. My qualitative analysis of these reviews reveals an ongoing discrepancy between narrative reviews and the rather limited one to five star review format. Interpretation of ratings into reviews, and vice versa, by review creators and consumers is neither straightforward nor is it objective.

Studying the effect of the textual comments in a content analysis of over 10,000 of eBay's online auction marketplace reviews, Pavlou and Dimoka (2006) found that "buyers read and take into consideration feedback text comments to compensate for the inability of numerical ratings to offer detailed information" (408). Single number ratings, it seems, are not particularly informative in and of themselves. Although it may seem efficient to boil down a variety of product attributes and their evaluations into a single number, this is a retailer practice that deserves careful scrutiny in our future research.

We must also question assumptions that ask us to reason from the general to the particular by taking the "average user rating" of a decontextualized 'average consumer' and then assert that we can discover something relevant that might apply to the actual world of

differentiated consumers with manifold needs. Instead, we should let the different needs, practices, perspectives, and experiences of our variegated world of consumers dictate the terms of our studies rather that any prefiguring presuppositions of objective desires, informational simplicity, or product utility. Developing the underpinnings of these presuppositions further in the next section, I deploy notions of ecological validity and contextual integrity to explore how to unite different approaches to researching online reviews and ratings.

External Validity and Contextual Integrity

In this concluding section, I wish to take a brief detour through the philosophy of science in order to discuss the tension between decontextualization and rigor that underpins much of the preceding sections. It is crucial to note, first, that science works through abstraction. Whether you conduct research as an economist, a psychologist, or an anthropologist, you take real world events, such as people using Internet sites to buy products, and then abstract out from the multitude of potential variables a greatly reduced set to explore and explain. We reduce the complexity of the world so that we can study it. We turn reality into relationships between constructs; this is how theory is built.

Theories allow us to see the forest for the trees. In many cases, they allow us to see the various types of trees that comprise the forest and, often, afford us insights into the relationship between them. No matter our approach, however, a certain risk of methodological reductionism always exists when we assume that we can understand the forest, or the phenomenon. A key issue is whether our mental maps of reality represent it well enough to be useful. Lynch (1982) points out that researchers must be especially careful about the external validity of theoretical

tests when there are unmanipulated background variables that might interact with the manipulated independent ones. Simonson (2015, 29) contends that the current Internet environment presents us with an environment that connects many previously disconnected elements, such as the ratings and reviews of millions of other consumer. Thus what previously was noise may now be signal.

Favoring external validity over the internal variety, future research into this area should carefully examine actual occurrences of phenomena in order to guide the selection or creation of variables, data, constructs, and relations. Such an approach fits well with Simonson's (2015, 29) advice that "studies that excel on the external validity dimension should be allowed to meet lower (within a reasonable range) internal validity standards". It is nearly impossible to attain depth of understanding as well as representativeness, generalizability, internal validity, and the definitive ruling out of rival explanations in single experimental or cultural theory studies. "But the field, as represented by journal editors and reviewers, may be better off being more lenient when evaluating studies that present potentially important findings, even if it is impossible to rule out some rival accounts" (Simonson 2015, 29).

When we engage with actual consumers, their needs, real marketing event and sites of consumption and communication, researchers in ostensibly applied fields such as consumer and marketing research must struggle mightily with questions of external validity. In the service of rigorous method or parsimonious theorizing, we often abstract away considerable context. The question is, do we abstract so much of it away that we do violence to our ability to understand the actual phenomenon? I term this notion of balance between context and abstract theory "contextual integrity" and argue that the decision to theorize at a particular level is a very important, but largely unmentioned, aspect of social science discipline. To spur the conversation,

13

Figure 4 presents illustrates the idea of a "Goldilocks approach" to matching theory with context. In that figure, we can see how adequate theorizing is a balancing act. Appropriate theory building occurs at a level of analysis poised between description and abstraction, between being overly complex and oversimplification, between being so close that it resembles a journalistic description and so distant that it is unrealistic and unrecognizable. The Goldilocks equilibrium lies in the middle and is, like the story's temperate bowl of porridge, just right.

Figure 4 About Here

FIGURE 4: FINDING THE RIGHT THEORETICAL BALANCE BETWEEN CONTEXT AND ABSTRACTION
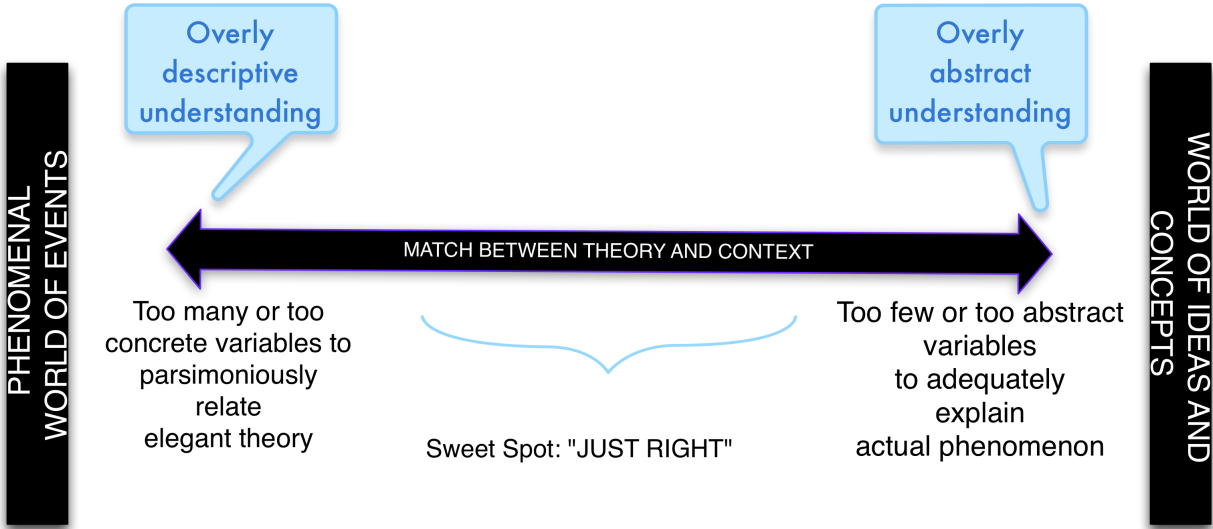
How does Figure 4 apply to the future study of online consumer-generated reviews and ratings? First it draws our attention to the appropriateness of the match between theory, context, method, and analysis. In some cases, for some decisions, simple and elegant parsimonious measures and highly abstract theory-driven approaches may be sufficient for understanding. However, in the current ever-expanding universe of mobile and stationary Internet communications, which links everyone to everything, enabling two billion very different

14

consumers to create, share, and consume a vast variety of social and emotional information, we face a far more complex consumption environment than ever before.

In this paper, I have argued that the world of online reviews and ratings reveals customers with diverse needs and subject positions, whose narrative reviews may not agree with their numerical ratings. I have discussed consumers whose subjective guidelines determine the resemblance of particular reviewers to themselves so that they can determine from their review whether a particular product will actually work for them. In addition, I have shown how rankings and ratings on Amazon.com present consumers with a complex social communication environment, a conversation that grants them a sense of social presence as well as opportunities for a range of expressive and explanatory options. I have also noted how the review and ratings environment on Amazon, or any other site, is only one part of a wide range of options in contemporary consumers' current communicational repertoire.

Attention to appropriate fit between theory and context leads us to question the use of terms such as "objective". It suggests that we not presume that one particular use of reviews or ratings is the correct or best one. It leads us to seek a deeper understanding of the multifaceted social, emotional, and relational properties of brands as they exist and expand in online reviews and ratings. It warns us to avoid using terms such as rational when describing consumers or their decision-making processes. Finally, it inspires us to initiate and value studies with high external validity, even if that means sacrificing some internal validity.

Simonson (2016) asserts "the growing prevalence and impact of user reviews was not anticipated" by consumer researchers investigating the early years of the Internet. However, in 1999, I wrote that online consumers "create reviews of products, giving informed, justified 'thumbs up' or 'thumbs down' evaluations of [products and services and, by recognizing this,]

15

marketers can have wide-ranging effects that inform and mediate consumer demand and consumption meanings across large numbers of others" (Kozinets 1999, 259-260). Many cultural and psychological consumer researchers such as Alladi Venkatesh, Russ Belk, Hope Schau, Janice Denegri-Knott, Mike Molesworth, Daiane Scaraboto, Marie-Agnes Parmentier, Nick Lee, Anne Schlosser, Wendy Moe, Donna Hoffman, Tom Novak, and Jonah Berger, have been studying and developing a range of sophisticated, contextualized theories to help explain how consumers create, share, and use online ratings and reviews. Our work should not be ignored or dismissed by BDT researchers, just as theirs should not be ignored or dismissed by us. De Langhe et al. (2015) perform a valuable service by following up on Simonson's (2015) call for more BDT work on the new informational environment, and by demonstrating how the Internet's effects on consumers cannot be easily explained as an increasing of their rationality or their resistance to branding. In the future, increased interaction and exchange of ideas between scholars of all disciplinary subfields working in this substantive area will undoubtedly be healthy for the continued growth and development of this burgeoning area of investigation.

REFERENCES

Arnould, Eric J. and Craig J. Thompson (2005), "Consumer Culture Theory: Twenty Years of Research," *Journal of Consumer Research*, 31(4), 868–893.

De Langhe, Bart, Philip M. Fernbach and Donald R. Lichtenstein (2016), "Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings," *Journal of Consumer Research*, XX (February).

Klaus, Philipp (2013), "The Case of Amazon.com: Towards a Conceptual Framework of Online Customer Service Experience (OCSE) Using the Emerging Consensus Technique (ECT)," *Journal of Services Marketing*, 27 (6), 443-457.

Kozinets, Robert V. (1997) "'I Want To Believe': A Netnography of The X-Philes' Subculture of Consumption," *Advances in Consumer Research*, Volume 24, ed., Merrie Brucks and Deborah J. MacInnis, Provo, UT: Association for Consumer Research, 470-475.

_____ (1999), "E-Tribalized Marketing? The Strategic Implications of Virtual Communities of Consumption," *European Management Journal*, 17 (3), 252-264.

_____(2002), "The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities," *Journal of Marketing Research*, 39 (February), 61-72.

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

Lynch Jr., John G. (1982), "On the External Validity of Experiments in Consumer Research," *Journal of Consumer Research*, 9 (December), 225–239.

Madianou, Mirca and Daniel Miller (2012), "Polymedia: Towards a New Theory of Digital

      Media in Interpersonal Communication," *International Journal of Cultural Studies*, 16

      (August), 169-187.

Pavlou, Paul A. and Angelika Dimoka (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), 392-414.

Simonson, Itamar (2015), "Mission (Largely) Accomplished: What's Next for Consumer BDT-

      JDM Researchers," *Journal of Marketing Behavior*, 1-32.

Simonson, Itamar and Emanuel Rosen (2014), *Absolute Value*, New York: Harper Collins.