

Running head: EVALUATING REVIEWS AND REVIEWERS ON YELP.COM

SAMPLE PUBLICATION 4

Evaluating Reviews and Reviewers on Yelp.com

Li Xiong and Andrea B. Hollingshead

University of Southern California

Presented at ICA 2010 Singapore

Evaluating Reviews and Reviewers on Yelp.com

Online communities are becoming important sources of information for consumers' purchase decisions. For example, many people share and consult local business and restaurant reviews on Yelp.com—a U.S. based website (McNeil, 2008). Unlike a traditional information source such as news media, it used to be difficult to assess the authenticity and usefulness of such information due to anonymity (Spears, Lea, & Postmes, 2006). For online communities like Yelp.com, user profiles may include both social demographic information such as names and locations, and reputation information that is related to a user's experience, expertise, popularity and quality of contributions. Such profile systems make it easier for individuals to verify the identities and evaluate the contributions of others (Ma & Agarwal, 2007). The present study investigated how the interaction of reputation and demographic information about reviewers influences consumers' evaluation of reviewers and reviews on Yelp.com.

The social psychological research on persuasion suggests that the influence of a message on individuals is determined by characteristics of both the source and content of the message (Hovland, Janus, & Kelly, 1953). People give different weight to the source and content of a message in situations that demand different levels of cognitive input and personal involvement (Ferran & Watts, 2008). In online communities like Yelp.com, restaurant reviews can include other information such as the profiles of the reviewers and the reactions to the reviews by other consumers. Because there are relatively high levels of information and relatively low levels of personal involvement in reading restaurant reviews, we expect that consumers will engage in heuristic processing. Thus, they will be influenced by characteristics of the reviewer in evaluating the quality of the content of the review (Chaiken, 1980; Chaiken, Liberman, & Eagly, 1989). To test this general hypothesis, in the present experiment, we systematically varied

characteristics of the reviewers and examined the impact on participants' evaluations of the review and the reviewer, while holding the content of the review constant.

Experience

Status is “one characteristic around which differences in cognitions and evaluations of individuals or social types of them come to be organized” (Berger & Zelditch, 1977). Status Characteristics Theory suggests that individuals in high-status positions in groups have more opportunities to contribute to tasks, receive higher evaluations of their contributions and exert more influence over low-status members (Berger, Cohen, & Zelditch, 1972). One proxy for status and reputation on Yelp.com may be the amount of experience a reviewer has in participating in this online recommendation community – writing and posting reviews. For example, if a person posts many reviews about many different restaurants, he or she should have substantial experience and knowledge about restaurants and reviewing. The quantity of reviews may indicate his or her relative knowledge about restaurants. Previous research supports this view. There is a positive association between individuals' knowledge sharing activities and their perceived expertise (Stasser, Stewart, & Wittenbaum, 1995). Research on online knowledge-sharing communities has observed that the quantity of knowledge contribution to open source software projects can be positively related to the perceived expertise of the contributors (Poston & Speier, 2008). Following this logic, we hypothesize that:

Hypothesis 1a: Reviews by experienced reviewers who have posted many reviews will be rated as more authentic, useful and persuasive than reviewers who have posted few reviews.

Hypothesis 1b: More experienced reviewers will be rated as more trustworthy, knowledgeable and experienced than less experienced reviewers.

Gender

People may rely on stereotypes based on demographic characteristics (such as gender) to infer others' knowledge in particular domains (Clark & Marshall, 1981; Krauss & Fussell, 1991; Wegner, 1987). Although gender is just one of many social categories that invoke stereotypes about what people know, the persistent domination of gender over race, age, and other cues as the basis for categorization across a variety of contexts (for a review see Zemore, Fiske, & Kim, 2000). Recently, research has demonstrated a general and positive association between the concepts “expert” and “male” when compared to the concepts “expert” and “female” (Folkes, Wong, & Hollingshead, 2010). Men tend to have more social influence and hold advantages over women across many domains of society such as business and education (Carli, 2001). In addition, women are considered less knowledgeable than men on specific subjects such as science and technology (Steele & Ambady, 2006). Even for gender-neutral tasks, people still develop expectations about others' competence based on previous assumptions of gender (Foschi, Enns, & Lapointe, 2003). This phenomenon suggests that people evaluate information based on the gender of the information source as well as the specific association between gender stereotypes and particular knowledge domains. When people evaluate restaurant reviews, those written by males will be perceived more favorably than those written by females.

Hypothesis 2a: Reviews by males will be rated as more authentic, useful and persuasive than those by females.

Hypothesis 2b: Male reviewers will be rated as more trustworthy, knowledgeable and experienced than female reviewers.

We also expected a statistically significant interaction effect such that experience would make more of a difference for female reviewers than male reviewers. In the absence of any

information about the reviewers other than their gender, female reviewers will be rated less favorably than male reviewers. However, women can be accepted as experts when they demonstrate their expertise and/or provide sufficient background and credentials (cf. Foschi, 2000). Thus, low status women may have to provide more evidence of their knowledge to be accepted as an expert (Foschi, 2000). Following this logic, we hypothesize that:

Hypothesis 3a: The effect of the experience difference of reviewers on the evaluation of reviews will be moderated by the gender of reviewers. Thus, reviews by less experienced female reviewers will be evaluated less positively than reviews of less experienced male reviewers, but there will be no statistically significant differences in the evaluations of reviews written by experienced male and female reviewers.

Hypothesis 3b: The effect of the expertise difference of reviewers on the evaluation of reviewers will be moderated by the gender of reviewers. Thus, less experienced female reviewers will be evaluated less positively than less experienced male reviewers, but there will be no statistically significant differences in the evaluations of experienced male and female reviewers.

Method

Design

The study employed a 3 x 2 x 2 repeated measures design. The between-subject factors were the reviewer's experience level (high: over 200 reviews or low: under 10 reviews) and the gender of the reviewer. The within-subject factor was the review. Three actual reviews of local restaurants near the University of Southern California campus in Los Angeles, California were selected from Yelp.com as the stimuli for the study. The three reviews were pilot tested to ensure they had similar length and were gender neutral (i.e. it was impossible to tell whether the review

was written by a male or female), but with different levels of valence. The first two reviews were positive about the restaurants with a four-out-of-five rating, and the last review was negative with a one-out-of-five rating.

The original review profiles were removed and replaced with profiles that were manipulated on the gender and expertise status factors. Specifically, each review was attached with a profile that depicted either a male or female reviewer, and showed the status of either a very experienced or less experienced reviewer (see Appendix I for a sample review). To rule out the potential influence of the reviewer's physical attractiveness, we selected profile pictures that were rated as "average" attractiveness on HotOrNot.com, a website where users share and rate the physical attractiveness of photos. Four experimental conditions were created, with each condition containing all three reviews. In addition, a control condition was created for which no reviewer information was provided for the three reviews.

Participants

The participants were 120 students from the University of Southern California. They were invited to participate in a study about how people use online restaurant reviews for extra credit in a large undergraduate course. Participants were randomly assigned to the five conditions. There were approximately equal cell sizes across the conditions.

Procedure

Each participant was emailed a link to a web survey. They were first asked to provide information about their previous visits and experiences at each of the three restaurants. Then, the participants were asked to read three reviews in random order. After reading each review, they first answered the question "How likely will you visit this restaurant in the next month, having read this review?" on a five-point-scale, with 1 indicating "very unlikely" and 5 "very likely".

This item was the measure of purchase intent. Then, they evaluated authenticity, usefulness, and quality of the review, and the trustworthiness, knowledge and experience of the reviewer. Each of these qualities was assessed with a one-item measure. For example, the evaluation of the experience of the reviewer was assessed with the question “In your opinion, how experienced is this reviewer in writing reviews?” with 1 standing for “very inexperienced” and 5 “very experienced”. After each question, open-ended questions probed participants to share the rationale for each of their ratings. After providing information about gender, race and age, participants were debriefed and thanked for their participation.

Analysis

A mixed methods ANCOVA model was run to assess the main and interaction effects in the hypotheses. Previous experiences at each restaurant, previous use of Yelp.com along with other online reviews, and physical attractiveness of each reviewer were controlled for in the analysis as covariates.

Results

Descriptive statistics

Complete responses from 118 of the 120 participants were used for the analysis. Two of the responses were not completed and thus not included in the analysis. The participants were predominantly female ($n = 93$, 88.6%) with an average age of 19.9 years ($SD = 1.32$). Ethnically, 48.6% of the participants are White/Caucasian, 30.5% Asian/Pacific Islander, 2.9% Black/African American and 5.7% Hispanic/Latino(a).

Hypothesis testing

Hypothesis 1a was not supported. There is no statistically significant main effect of reviewer experience on participants' evaluation of the usefulness, persuasiveness and quality of

reviews. However, hypothesis 1b was partially supported. The experience of the reviewer, which was indicated by the number of reviews he or she has posted on Yelp.com, had a significant main effect on participants' evaluation of the experience of the reviewer, $F(1, 68) = .423, p = .04$. As shown in Table 1 (see Appendix 2), a reviewer who had posted more than 200 reviews on Yelp.com was deemed as more experienced than a reviewer who has posted only 5 reviews, $t(68) = 2.06, p = .04$, and more experienced than the control condition where no reviewer information was disclosed, $t(71) = 2.57, p = .01$. The main effect of experience on perceived trustworthiness of the reviewer was not statistically significant but in the expected direction, $F(1, 59) = 3.38, p = .07$. Unexpectedly, the reviewers in the control condition were considered to be more trustworthy than less experienced reviewers, $t(65) = 3.00, p = .004$. The experience indicator had no statistically significant effect on the perception of the knowledge of reviewers.

Hypotheses 2a and 2b were both partially supported. Table 2 shows statistically significant main effects of reviewer gender on participants' evaluation of the persuasiveness and usefulness of the reviews, and reviewers' knowledge, $F(1, 34) = 4.40, p = .04$, $F(1, 67) = 4.69, p = .03$, and $F(1, 80) = 7.91, p = .006$ respectively. Specifically, male reviewers were considered more knowledgeable about the restaurants than female reviewers, $t(80) = 2.81, p = .006$. Reviews by male reviewers were more persuasive than those by both female and anonymous reviewers, $t(34) = 2.09, p = .04$ and $t(29) = 2.48, p = .02$ respectively. In addition, reviews were considered more useful when the reviewers were male than female, $t(67) = 2.16, p = .03$. Unexpectedly, comparing the experimental conditions to the control condition where no information about reviewers was disclosed, the control condition was viewed as more

trustworthy than female reviewers, $t(67) = 2.87, p = .006$, and their reviews were also rated more authentic than a female reviewer, $t(75) = 2.12, p = .04$.

Hypothesis 3a and 3b were also partially supported. As shown in Table 3, there was a statistically significant interaction effect of reviewer gender and experience status on participants' evaluation of the persuasiveness of the reviews, and the trustworthiness and experience of the reviewers, $F(4, 30) = 2.77, p = .05$, $F(4, 52) = 3.54, p = .01$, and $F(4, 68) = 2.75, p = .04$ respectively. A male reviewer who had posted many reviews was considered as more persuasive, experienced and trustworthy than a female reviewer who had posted very few reviews, $t(31) = 2.82, p = .01$, $t(67) = 2.67, p = .01$, and $t(46) = 2.60, p = .01$ respectively. Experienced male reviewers were also deemed more experienced and persuasive than the no-information-about-reviewer control condition, $t(70) = 2.96, p = .004$ and $t(30) = 2.57, p = .02$ respectively.

In contrast, female reviewers had no such advantage: there was no statistically significant difference in participants' evaluation of persuasiveness, trustworthiness and experience between a female expert reviewer and a male or female reviewer who has posted very few reviews. Surprisingly, the no-information-about-reviewer control condition was deemed more trustworthy than either female or male reviewer with few reviews, $t(61) = 3.42, p = .001$ and $t(62) = 2.15, p = .04$ respectively. However, comparing the experimental conditions to the control condition, we found that reviews in the control condition were considered to be even more authentic than those by female reviewers with many reviews, $t(82) = 2.39, p = .02$.

Discussion

This study examined the influence of the attributes of reviewers on consumers' evaluation of reviewers in online communities. We investigated how people evaluate restaurant

reviews and reviewers on Yelp.com based on the experience and gender of reviewers. The preliminary results supported our general hypothesis that reviewer attributes can significantly influence how people evaluate and judge the content of online information. However, the results did not always support our initial hypotheses.

As predicted, our results showed that people evaluate restaurant reviews based on the gender and experience of reviewers. This effect of social attribution is well documented in previous research in the offline context (e.g. Chaiken, et al., 1989; Ridgeway & Diekema, 1992; Weisband, Schneider, & Connolly, 1995). In the context of Yelp.com, we first observed that males were considered as more knowledgeable, persuasive and useful sources of opinions about local restaurants than females. Given that our sample consisted predominantly of female participants, we cannot draw the conclusion that males are more influential sources of information than females in online communities. Yet, it is reasonable to argue that gender is an important social cue that influences the evaluation of online information. Furthermore, we observed that the number of reviews that a reviewer posted on Yelp.com was positively related to perceived experience in writing reviews. The association of male gender with expertise moderated the effect of experience status on the evaluation of reviews and reviewers.

However, our results did not support the idea that personal information can influence all aspects of information evaluation in the same way. On one hand, the number of reviews that a reviewer contributed on yelp.com did not affect consumers' perceptions of the knowledge of reviewers or the quality of reviews. Thus, experience was not a proxy for expertise or influence.

On the other hand, the disclosure of personal information did not appear to positively influence the evaluation of trustworthiness and authenticity of reviewers. On the contrary, the control condition where reviews had no reviewer information was perceived to be more

trustworthy and authentic than the female condition and the condition of reviewers with very few reviews. We speculate that in our context of Yelp.com, judgment of trustworthiness and authenticity might be separate and different from the process in which people evaluate the extrinsic value of reviews and reviewers. As suggested in the Heuristic-Systematic Model (Chaiken, et al., 1989) and the Elaboration Likelihood Model (Petty & Cacioppo, 1986), the context of information adoption might influence how people process social attributes of information sources as decision cues. People's objective for reading reviews may vary. Therefore their evaluation processes can be influenced by different cues.

Our study asked the participants to quickly evaluate reviews based on the personal attributes of reviewers. A post-hoc analysis on our results showed that on average, participants' intention to visit the restaurants was positively correlated with their evaluation of the quality, usefulness and positivity of the reviews, $r^2 = .52, p < .01$, $r^2 = .22, p < .05$, and $r^2 = .35, p < .01$ respectively. This revealed a utilitarian rather than socializing objective among our participants, which alluded to a potential contextual effect. Although the disclosure of personal information facilitated value judgment (usefulness, knowledge, experience, and persuasiveness), it did not facilitate trust, a social judgment that needs to be sustained and reinforced in involved interactions and extended relationships (Jarvenpaa, Shaw, & Staples, 2004; McKnight, Cummings, & Chervany, 1998).

Nevertheless, we still found it puzzling that the no-information about the reviewer control condition received higher ratings of reviewers' trustworthiness and authenticity than female reviewers. This point can be illustrated by a participant's comment in the female experienced reviewer condition, "There's no way of telling how credible the reviewer is...At times it's difficult to filter through information or make judgment calls on what information is actually

valuable to you though.” It appeared that personal information about the reviewer served as a heuristic cue, which influence information evaluation even though those social cues were not relevant.

This study examined how reviewer characteristics influence how consumers evaluate the content of reviews. For a given restaurant on yelp.com, there may be hundreds of reviews. People may use heuristic cues such as gender and reputation as a filter to select among the many opinions and reviews. However, the results suggested that reviewers’ demographic and social information influence information evaluation in unexpected ways.

References

- Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, 37, 241-555.
- Berger, J., & Zelditch, M. J. (1977). Status characteristics and social interactions: The status organizing process. In J. Berger, M. H. Fisek, R. Z. Norman & M. J. Zelditch (Eds.), *Status characteristics and social interaction*. New York, NY: Elsevier.
- Carli, L. (2001). Gender and Social Influence. *Journal of Social Issues*, 57(4), 725-741.
- Chaiken, S. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion *Journall of Personality and Social Psychology*, 39(5), 752-766.
- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 212-252). New York: Guilford Press.
- Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. Joshi, I. Sag & B. Webber (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge, UK: Cambridge University Press.
- Ferran, C., & Watts, S. (2008). Videoconferencing in the field: A heuristic processing model. *Management Science*, 54(9), 1565-1578.
- Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology*, 26(1), 21-42.
- Foschi, M., Enns, S., & Lapointe, V. (2003). Processing performance evaluations in homogeneous task groups: Feedback and gender effects. *Advances in Group Processes*, 18, 185-216.
- Hovland, C., Janus, I. J., & Kelly, H. H. (1953). *Communication and Persuasion: Psychological Studies of Opinion Change*. New Haven, CT: Yale University Press.
- Jarvenpaa, S. L., Shaw, T. R., & Staples, D. S. (2004). Toward contextualized theories of trust: The role of trust in global virtual teams. *Information Systems Research*, 15(3), 250-267.
- Krauss, R. M., & Fussell, S. R. (1991). Constructing shared communicative environments. In L. B. Resnick, J. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition*. Washington, DC: APA Press.
- Ma, M., & Agarwal, R. (2007). Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. *Information Systems Research*, 18(1), 42-67.
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *The Academy of Management Review*, 23(3), 473-490.
- McNeil, D. G. (2008, November 4). Eat and tell, *The New York Times*.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.
- Poston, R. S., & Speier, C. (2008). Knowledge management systems usage: Rating scheme validity and the effort-accuracy trade-off. *Journal of Organizational and End User Computing*, 20(1), 1-16.
- Ridgeway, C. L., & Diekema, D. (1992). Are gender differences status differences? . In C. L. Ridgeway (Ed.), *Gender, interaction, and inequality*. New York, NY: Springer-Verlag.

- Spears, R., Lea, M., & Postmes, T. (2006). Computer-mediated communication and social identity. In A. Joinson, K. Y. A. McKenna, T. Postmes & U. Reips (Eds.), *Oxford Handbook of Internet Psychology* (pp. 253-269). Oxford, UK: Oxford University Press.
- Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology, 1*, 244-265.
- Steele, J. R., & Ambady, N. (2006). "Math is Hard!" The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology, 42*(4), 428-436.
- Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly, 29*(1), 35-57.
- Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.
- Weisband, S. P., Schneider, S. K., & Connolly, T. (1995). Electronic communication and social information: Status salience and status differences. *Academy of Management Journal, 38*, 1124-1151.
- Zemore, S. E., Fiske, S. T., & Kim, H. (2000). Gender stereotypes and the dynamics of social interaction. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 207-242). Mahwah, NJ: Erlbaum.

Table 1

Comparison of Review Evaluation Across Experience Levels

	Experienced Reviewer	Inexperienced Reviewer	Control
How experienced is this individual in writing reviews?	3.83	3.49	4.53
How much do you trust the opinion of the reviewer?	3.99	3.86	4.47

Note. 1 = lowest rating, 5 = highest rating

Table 2

Comparison of Review Evaluation Across Gender Levels

	Male	Female	Control
How likely will you visit the restaurant in the next month?	3.92	3.59	3.36
How would you rate the usefulness of this review?	3.91	3.52	3.82
How knowledgeable about restaurants is this reviewer?	3.69	3.19	3.21
How much do you trust the opinion of the reviewer?	3.83	3.49	4.53
How authentic does this review seem?	3.99	3.86	4.47

Note. 1 = lowest rating, 5 = highest rating

Table 3

Comparison of Review Evaluation Across Gender X Experience Levels

	Male Experienced Reviewer	Female Experienced Reviewer	Male Inexperienced Reviewer	Female Inexperienced Reviewer	Control
How likely will you visit the restaurant in the next month?	4.00	3.76	3.84	3.41	3.36
How much do you trust the opinion of this reviewer?	3.98	3.76	3.68	3.22	4.53
How experienced is this reviewer in writing reviews?	3.93	3.50	3.41	3.20	2.97
How authentic is this review?	4.11	3.68	3.87	4.05	4.47

Note. 1 = lowest rating, 5 = highest rating

Appendix 1

Sample Review for the Condition of Female X Experienced Reviewer

The Coffee Bean & Tea Leaf
 Neighborhood: South Los Angeles
 3726 S Figueroa St
 Los Angeles, CA 90007(213) 745-4963



Alex
Reviews posted
by this user: 249

7/27/2009

This is a unique Coffee Bean... I don't know if it's one of a kind, but it definitely is suited to serve the student community at USC.

The main obvious difference from other locations is that they serve light breakfast... build-your-own cereal, oatmeal, and yogurt served all day.

Prices are kinda high (I wouldn't pay \$3 for a scoop of Fruitty Pebbles) but I guess it's the premium for convenience.

Inside, there is a plenty of seating.. but surprisingly (or not) I see NO outlets anywhere... perhaps it's to keep students from camping out? Either way it's annoying to me.

The outside patio is the best feature of this store... it's huge (I count at least 25 of those bistro tables), complete with a fountain and firepit. It's also enclosed by an iron/brick fence, so "intrusions" from the sidewalk are minimized.

Not much to say about the staff... they were nice and friendly enough to me, and made my drink well. I find that a lot of times, your own attitude matters just as much as the service staff!