

Association for Information Systems AIS Electronic Library (AISeL)

ICIS 2003 Proceedings

International Conference on Information Systems
(ICIS)

December 2003

Effects of Information Presentation on Perceived Reputation in Virtual Communities: A Controlled Experiment

Kevin Kuan
University of Michigan

Judith Olson
University of Michigan

Richard Gonzales
University of Michigan

Follow this and additional works at: <http://aisel.aisnet.org/icis2003>

Recommended Citation

Kuan, Kevin; Olson, Judith; and Gonzales, Richard, "Effects of Information Presentation on Perceived Reputation in Virtual Communities: A Controlled Experiment" (2003). *ICIS 2003 Proceedings*. 91.
<http://aisel.aisnet.org/icis2003/91>

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2003 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

EFFECTS OF INFORMATION PRESENTATION ON PERCEIVED REPUTATION IN VIRTUAL COMMUNITIES: A CONTROLLED EXPERIMENT

Kevin K. Y. Kuan
Business School
University of Michigan
Ann Arbor, MI USA
kkuan@umich.edu

Judith S. Olson
School of Information
University of Michigan
Ann Arbor, MI USA
jsolson@umich.edu

Richard Gonzalez
Department of Psychology
University of Michigan
Ann Arbor, MI USA
gonzo@umich.edu

Abstract

Reputation systems have become increasingly popular in virtual communities as a way to record and communicate the reputation information of the members. However, different reputation systems use different presentation formats and their effects on decisions in terms of evaluating positive and negative ratings remain unclear. A controlled experiment is proposed using the preference ladder procedure to elicit subjective preferences in three commonly used presentation formats. One format presents the negative and positive ratings side by side; one presents the information as a percentage of total ratings that are positive; the third presents the difference between the positive and negative ratings, the format used by eBay. Results of the preliminary data analysis suggest that people weigh the positive and negative information to different extents in the three formats. Presenting reputation in the difference format tends to make a person weigh the negative information less, making the person more forgivable. The finding is possibly due to the salience of the negative ratings in the various presentations.

Keywords: Information presentation, reputation systems, trust, virtual communities

Introduction

The Internet has provided a global virtual community that enables individuals to interact and establish relationships with other individuals independent of location. However, individual-to-individual relationships in virtual community often involve higher uncertainty, to some extent due to the loss of social cues over the communication medium, and therefore require higher trust (Olson and Olson 2000). As a Peter Steiner's famous cartoon in *The New Yorker* (July 5, 1993, p. 61) illustrates, "On the Internet, nobody knows you're a dog."

Research has told us that "trust needs touch" (Olson and Olson 2000). In the physical community, trust can be built by repeated face-to-face interactions. In the virtual community, face-to-face interactions are almost impossible, and repeated interactions are

often limited to business-to-consumer markets. For consumer-to-consumer markets such as eBay, buyers are unlikely to have repeated transactions with the same sellers. As a result, effective communication of trustworthiness is needed to sustain virtual markets.

A commonly used mechanism for communicating trustworthiness is reputation. Reputation provides information on the person's likelihood of carrying out certain behaviors in the future (Whitmeyer 2000; Wilson 1985). In the physical community, reputation is often communicated informally through repeated interactions or word of mouth such as gossiping. While word-of-mouth may be effective in small, close communities, it becomes ineffective as the community grows. As a result, documenting reputation formally using a reputation system has become increasingly popular in online communities. Members can rely on information provided by other members to reduce the risk (Lee 1998).

Reputation systems can be generally categorized based on the mechanism under which reputation information is collected. One mechanism, which is more common in business-to-consumer markets such as TRUSTe and the Better Business Bureau, is based on reviews by reputable authorities. Another mechanism, which is more common in commerce-to-commerce markets, is based on peer reviews. Examples of peer-reviewed reputation systems include Download.com and eBay Feedback Forum. Online reputation systems are emerging as an important mechanism for quality signaling and as a quality control mechanism in online trading communities (Dellarocas 2001; Kollok 1999; Resnick et al. 2002). Reputation systems are generally believed to be effective in building trust among the potential trading parties in an online market (Walden 2000), and to improve the effectiveness of the market (Adler 2001).

Different reputation systems present reputation information differently. For example, eBay Feedback Forum is a binary rating system under which each user buyer or seller can be rated as positive (+1) or negative (-1). The overall rating is represented as the *difference* between the number of positive ratings and negative ratings. Download.com uses a similar binary rating system under which software can be rated as “thumbs up” (good) or “thumbs down” (bad). However, the overall rating of the software is represented as the *percentage* of ratings that are positive. While the two presentation formats are equivalent in terms of the content communicated, it is unclear whether people receive and use the information the same way under different formats. We believe that the usefulness of a reputation system depends not only on the quality of the information captured by the system but also the quality of the information presented to the user. By understanding how users interpret the underlying information provided by a reputation system, system designers can improve the presentation of information so that it can be easily and accurately interpreted by the users. Therefore, this study investigates how different presentation of the same piece of information regarding one's reputation may be evaluated differently under different formats.

Literature Review

Reputation information provides information regarding the probability (likelihood) of positive outcomes (favorable behavior of the trustee) and negative outcomes (unfavorable behavior of the trustee) related to the trust decision (Whitmeyer 2000). It has been argued that risky decisions that require trust are generally based on calculations of the trustee's costs and benefits of cooperating versus cheating as evaluated by the trustor (Ba and Pavlou 2001; Dasgupta 1988; Hart and Saunders 1998; Williamson 1993).

Positive and negative ratings are generally used to estimate the likelihood of gaining or losing (i.e., positive ratings are associated with possible gains), and negative ratings are associated with possible losses. However, losses were generally found to have greater impact than gains, an asymmetry well known in decision making (Coombs and Lehner 1984). Specifically, prospect theory states that separate weights are assigned gains and losses, people are more sensitive and averse to loss than gain (Kahneman and Tversky 1979). Therefore, negative feedback was found to have a greater impact than positive feedback (Johnson et al. 1988; Pyszczynski and Greenberg 1981; Roese 1997; Wong and Weiner 1981).

However, it has been pointed out in the literature that decision making under risk and uncertainty is not always based on full calculations of costs and benefits. Rather some simple heuristics are used and people are sensitive to the framing of the information provided (Tversky and Kahnman 1974). Different reputation systems present reputation ratings in different numerical formats. For example, reputation systems used by eBay and Yahoo! Auction use the frequency format, while those by Amazon.com and Yahoo! Shopping use the decimal format. Is a person with 95 positive ratings and 5 negative ratings viewed the same when his ratings are presented as “95 positive, 5 negative,” “95% positive out of 100 ratings,” or “90 net positive out of 100 ratings”?

It is often assumed that a rational decision maker interprets the values under different numerical formats equivalently, sometimes referred as the *numerical equivalence assumption* (Cohen et al. 2002). Nevertheless, people do interpret numerical formats differently, with different numerical formats associated with different abstract representations of amount. The implication that different representations of a common abstract structure can cause dramatically different cognitive behaviors is sometimes referred to as *representation determinism* (Zhang and Norman 1994). Representation determinism implies that the form of external representations determine what information can be perceived, what processes can be activated, and what structures can be discovered from the specific representation (Zhang 1997). In fact, it is known that people generally process information in the fashion related to the way it is displayed (Bettman and Kakkar 1977). This suggests that presenting feedback as the *difference* between the number of positive and negative ratings (referred to as the *difference format*) is unlikely to be evaluated the same way as presenting feedback as a *percentage* of positive ratings to the total number of ratings (referred to as the *percentage format*), or simply presenting the positive and negative ratings *separately* (referred to as the *separate format*). Based on the literature review, we have the following hypotheses:

H1: Negative ratings have a stronger effect than positive ratings regardless of the presentation format.

The hypothesis implies that an additional negative rating requires more than one positive rating to be compensated. In other words, person A with 95 positive ratings and 5 negative ratings should always be perceived as less-reputable than person B with 90 positive ratings and 0 negative.

H2: Presentation format moderates the effects of negative ratings on perceived reputation. Negative ratings have the strongest effect under the separate format, followed by the percentage format, and then the difference format.

The hypothesis implies that a person with more negative ratings than another person would require a different number of additional positive ratings under different presentation formats in order to make the two people be perceived as equally reputable. This can be explained by the fact that negative ratings have stronger effects than positive ratings, and that negative ratings are the most salient under the separate format, followed by the percentage format and then the difference format.

Experimental Design

This study looks at how different representations of the same piece of information in different formats affect the subjective evaluation of someone's reputation. Three representation formats will be evaluated. The first format, referred to as the separate format, presents overall reputation by simply displaying the number of positive and negative ratings separately. The second format, referred to as the difference format, presents overall reputation as the difference between the number of positive and negative ratings, together with the total number of ratings. This format is currently used by Websites such as eBay and Yahoo. The last format, referred to as the percentage format, presents overall reputation as a percentage of the number of positive ratings over the total of ratings, together with the total number of ratings. This format is currently used by Websites such as Download.com.

A one-hour experiment is being conducted with 78 subjects to-date making a series of 40 rounds of decisions based on a situation commonly used in trust and reputation research (Berg et al. 1995). In each round, the subjects are required to indicate their preferences between two hypothetical borrowers to which the subjects would lend \$1,000 based on the borrowers' credit ratings (reputation ratings). The credit ratings of the two borrowers vary across the total number of ratings (50, 100, 200, and 400), and the percentages of negative ratings (0%, 4%, 8%, 16%, and 32%) at each level of total number of ratings. In other words, the 40 rounds consist of 4 levels of total number of ratings, with 10 possible pairs of different ratings at each level, resulting in 40 different pairs of credit ratings. To reduce fatigue, subjects were given a 10 minute break after round 20. Also, the 40 pairs of credit ratings were assigned to the subjects in random order, with each pair assigned to some subjects in the earlier rounds, and some subjects in the later rounds.

In each round, the preferences on the two borrowers are elicited using the *preference ladder* procedure (Gonzalez and Wu 1999). An example of a preference ladder is provided in Table 1. Each ladder consists of 10 rungs with the credit rating of borrower A at a particular level of total number of ratings (e.g. 100, plus a small number), and a particular percentage of negative rating (e.g., 4%). The credit rating of borrower B in rung 1 has a higher percentage of negative ratings (e.g., 16%), but also a higher number of positive ratings. However, the *difference* between the positive and negative ratings of borrower A ($102 - 4 = 98$) is equal to the *difference* between the positive and negative ratings of borrower B ($114 - 16 = 98$) at rung 1. This allows us to evaluate

whether accounting overall reputation as the difference between positive and negative ratings is a reasonable representation. If that is the case, the two borrowers should be perceived as equally reputable at rung 1 (i.e., average rating at rung 1 should be 3). In the following nine rungs, the credit rating of borrower A remains constant, while the credit rating of borrower B improves with a constant number of positive ratings added. The preference ladder also allows us to evaluate the representativeness of accounting overall reputation as the percentage of positive ratings over total ratings. If that is the case, the two borrowers should be perceived as equally reputable at the rung where the two borrowers have the same reputation ratings in terms of percentage.

Table 1. Example of Preference Ladder under the Separate Format Condition

Rung	Credit Rating of Borrower A	Clearly Prefer A	Slightly Prefer A	About the Same	Slightly Prefer B	Clearly Prefer B	Credit Rating of Borrower B
1	102 positive 4 negative	X					114 positive 16 negative
2	102 positive 4 negative	X					138 positive 16 negative
3	102 positive 4 negative		X				162 positive 16 negative
4	102 positive 4 negative		X				186 positive 16 negative
5	102 positive 4 negative		X				210 positive 16 negative
6	102 positive 4 negative			X			234 positive 16 negative
7	102 positive 4 negative				X		258 positive 16 negative
8	102 positive 4 negative				X		282 positive 16 negative
9	102 positive 4 negative				X		306 positive 16 negative
10	102 positive 4 negative					X	330 positive 16 negative

Preliminary Results

More analyses will be conducted as more data is collected. Nevertheless, the preliminary results tend to support the hypotheses. Figure 1 provides an example in which the average ratings of the two borrowers over the 10 rungs when borrower A starts with 2 percent negative rating and borrower B starts with 8 percent negative ratings. As hypothesized, a negative rating has a stronger effect than a positive rating. This is demonstrated by the fact that although the difference between the number of positive and negative ratings is the same for borrower A ($102 - 4 = 98$) and borrower B ($114 - 16 = 98$) at rung 1, borrower A is strongly preferred to borrower B for all three conditions.

However, the number of additional positive ratings required to make borrower B be perceived as being as reputable as borrower A is different under different presentation formats. This is reflected by the different cutoff points at which borrower B and borrower A are equally preferred. Negative effects have the weakest effect under the difference format, as shown in the figure that borrower A is equally preferred to borrower B at about rung 6. Under the percentage format, person A is equally preferred to person B at about rung 7. In other words, borrower A with 102 positive and 4 negative ratings is perceived as equally reputable as borrower B with 234 positive and 16 negative ratings when the ratings are presented in the difference format. However, borrower B would need 258 positive and 16 negative ratings in order to be perceived as equally reputable as borrower A when the ratings are presented in the percentage format. Under the separate format, borrower B would need 330 positive and 16 negative ratings.

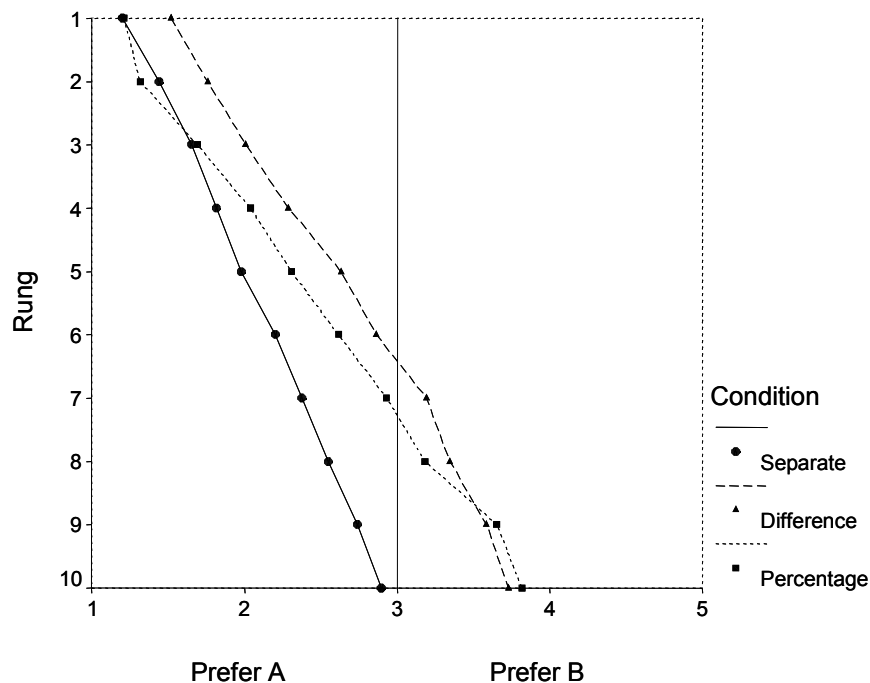


Figure 1. Average Ratings of the Two Borrowers under the Three Presentation Formats in One of the Preference Ladders

The results suggest that negative ratings have different effects under different presentation formats, which possibly can be explained by the salience of negative ratings under each format. The effects of negative ratings under different presentation formats may induce difference in members' behavior in a virtual community. For example, a member of a virtual community presenting reputation using the separate format would be less forgivable for his/her negative behavior since a negative rating requires more additional positive ratings to compensate. As a result, a member may be more conscientious, wanting to avoid negative ratings such as signing up under a new alias.

Acknowledgments

The work presented here was supported in part by NSF Grant IIS 9977923 and IIS-0233279.

References

Adler, P. S. "Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism," *Organization Science* (12:2), 2001, pp. 215-234.

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

Berg, J., Dickhaut, J., and McCabe, K. "Trust, Reciprocity, and Social-History," *Games and Economic Behavior* (10:1), 1995, pp. 122-142.

Bettman, J. R., and Kakkar, P. "Effects of Information Presentation Format on Consumer Information Acquisition Strategies," *Journal of Consumer Research* (3:4), 1977, pp. 233-240.

Cohen, D. J., Ferrell, J. M., and Johnson, N. "What Very Small Numbers Mean," *Journal of Experimental Psychology-General* (131:3), 2002, pp. 424-442.

Coombs, C. H., and Lehner, P. E. "Conjoint Design and Analysis of the Bilinear Model: An Application to Judgments of Risk," *Journal of Mathematical Psychology* (28:1), 1984, pp. 1-42.

Dasgupta, P. "Trust as a Commodity," in *Trust: Making and Breaking Cooperative Relations*, D. Gambetta (ed.), Basil Blackwell, Inc., Oxford, 1988, pp. 49-72.

- Dellarocas, C. "Analyzing the Economic Efficiency of eBay-like Online Reputation Reporting Mechanisms," in *Proceedings of the 3rd ACM Conference on Electronic Commerce*, Tampa, FL, 2001, pp. 171-179
- Gonzalez, R., and Wu, G. "On the Shape of the Probability Weighting Function," *Cognitive Psychology* (38:1), 1999, pp. 129-166.
- Hart, P., and Saunders, C. "Emerging Electronic Partnerships: Antecedents and Dimensions of EDI Use from Supplier's Perspective," *Journal of Management Information Systems* (44:4), 1998, pp. 87-111.
- Johnson, E. J., Payne, J. W., and Bettman, J. R. "Information Displays and Preference Reversals," *Organizational Behavior and Human Decision Processes* (42:1), 1988, pp. 1-21.
- Kahneman, D., and Tversky, A. "Prospect Theory: Analysis of Decision under Risk," *Econometrica* (47:2), 1979, pp. 263-291.
- Kollock, P. "The Production of Trust in Online Markets," in *Advances in Group Processes*, E. J. Lawler, M. Macy, S. Thyne, and H. A. Walker (eds.), JAI Press, Greenwich, CT, 1999.
- Lee, H. G. "Do Electronic Marketplaces Lower the Price of Goods?," *Communications of the ACM* (41:1), 1998, pp. 73-80.
- Olson, J. S., and Olson, G. M. "Interpersonal Trust on the Internet: i2i in E-Commerce," *Communications of the ACM* (43:12), 2000, pp. 41-44.
- Pyszczynski, T. A., and Greenberg, J. "Role of Disconfirmed Expectancies in the Instigation of Attributional Processing," *Journal of Personality and Social Psychology* (40:1), 1981, pp. 31-38.
- Resnick, P., and Zeckhauser, R. "Trust among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," in *Advances in Applied Microeconomics*, M. R. Baye (Ed.), Elsevier Science, Amsterdam 2002, pp. 127-157.
- Roese, N. J. "Counterfactual Thinking," *Psychological Bulletin* (121), 1997, pp. 133-148.
- Tversky, A., and Kahneman, D. "Judgment under Uncertainty: Heuristics and Biases," *Science* (185:4157), 1974, pp. 1124-1131.
- Walden, E. "Some Value Propositions of Online Communities," *Electronic Markets* (10:4), 2000 (available online at <http://www.electronicmarkets.org/modules/pub/view.php/electronicmarkets-70>).
- Whitmeyer, J. M. "Effects of Positive Reputation Systems," *Social Science Research* (29:2), 2000, pp. 188-207.
- Williamson, O. E. "Calculativeness, Trust, and Economic Organization," *Journal of Law and Economic* (36:1), 1993, pp. 453-486.
- Wilson, R. "Reputations in Games and Markets," in *Game-Theoretic Models of Bargaining*, A. E. Roth (ed.), Cambridge University Press, Cambridge, New York, 1985, 27-62.
- Wong, P. T. P., and Weiner, B. "When People Ask Why Questions, and the Heuristics of Attributional Search," *Journal of Personality and Social Psychology* (40:4), 1981, pp. 650-663.
- Zhang, J. "The Nature of External Representations in Problem Solving," *Cognitive Science* (21), 1997, pp. 179-217.
- Zhang, J., and Norman, D. A. "Representations in Distributed Cognitive Tasks," *Cognitive Science* (18), 1994, pp. 87-122.