ABSTRACT

Nowadays, many of sellers who retail products in C2C online marketplace improve their evaluations to systematically engage in evaluation manipulation behaviors. The purpose of this study not only examines the influence of sellers’ manipulation behaviors on percentage of positive ratings and comment orientation, but also confirm whether buyers’ experience moderating the relationship of sellers’ manipulation and buyers’ evaluations or not in evaluation system. The results show not only sellers’ manipulation behaviors influence on percentage of positive ratings, but also the interaction between sellers’ manipulation and buyers’ experience impacts on both comment orientation and percentage of positive ratings while the interaction leads to significant differences among the three types of seller manipulation.

KEYWORDS: Online evaluation system, Rating score, Comment orientation, Manipulation, Buyer experience

INTRODUCTION

With the rapid growth of e-commerce, consumers are becoming accustomed to online shopping. According to a global online shopping survey, about 93% of consumers have purchased a product on Internet (Pitney Bowes Inc., 2011). Meanwhile, about 69% of consumers are reported to be highly capable of making online purchases (Maxwell & Hudson, 2011). Internet provides a convenience channel to customers for online shopping anytime and anywhere. Nowadays, consumers buy virtually anything we can imagine from online marketplace through only a few clicks/taps. Nevertheless, often times the credibility of sellers is still a concern, especially in an online marketplace where smaller sellers gather. Insufficient trust between sellers and buyers on online marketplace interferes with consumers’ transaction decisions (Utz, Matzat, & Snijders, 2009). Therefore, most online marketplaces adopt feedback mechanisms, such as a customer evaluation system, to promote mutual trust and enhance information transparency between sellers and buyers (Bolton et al., 2004).

Online shoppers tend to balance the information asymmetry in relation to sellers by first referencing the consumer evaluation system readily provided on the online marketplace (Ba & Pavlou, 2002; Zhou and Windle, 2008). Some would also seek comments and suggests external to the online marketplace. However, the limitations of evaluation systems are noted, which includes anonymity, equitableness, online identity and online fraud (Resnick et al., 2000).
For instance, sellers could intentionally price a few items unusually cheap to attract buyers in order to accumulate positive ratings; some sellers even negotiate with buyers to not leave negative ratings, while some sellers retaliate against buyers who give negative ratings (Gregg & Scott, 2006). Steiner (2003) confirms that about 19% of eBay users have received retaliatory feedback and about 16% have been a victim of feedback extortion. On the other hand, to stimulating positive ratings, some sellers engage in systematic manipulations by constantly posting reminders, attempting to influence buyers’ evaluation and to sustain good image. It is plausible that some buyers would not leave negative feedback on evaluation systems even if they are unhappy with the sellers, to avoiding receiving threats from sellers.

Today, online marketplaces such as eBay.com and Amazon.com, which have earned world-wide reputation, seem to be consumers’ first choice when they purchase online. Yet in Asia the choices may not be the same. For example, C2C online marketplace Taobao.com has become one of the best known online shopping websites in Asia, and it accounts for over 80% of the e-commerce transactions in China, according to the company overview of Alibaba group (2013). Taobao.com was launched in 2003 by the Alibaba Internet business group which was founded in 1999. In 2014, there were approximately 500 million registered users on Taobao.com, which was visited by over 60 million unique guests per day (Taobao.com, 2014). Taobao.com provides a transaction platform for individual sellers and small businesses to sell retail items through auction or buy-it-now (BIN) option. With BIN, sellers set up fixed-price items and buyers purchase the items directly without going through an auction process (Ye et al., 2009). Like most online evaluation systems, Taobao.com’s consists of both rating scores and text comments. In addition, the evaluation is two-way, meaning not only buyers can evaluate sellers, sellers can also evaluate buyers. A composition of smaller sellers is typical to online marketplaces, thus external information on sellers are scarce and more difficult to acquire. Thus, when making purchases on a C2C online marketplace, buyers almost solely rely on the evaluation system provided by the online marketplace in making their buying decisions (Huang et al., 2011; Song & Baker, 2007). Sellers have to take the rating scores and text comments they received seriously, because they provide clear and direct credibility evidences for consumers.

Many previous studies have investigated various issues relating to online evaluation systems in online marketplaces. For example, Bolton et al. (2004) explored the effectiveness of electronic reputation mechanisms, Resnick et al. (2000) discussed the relationship between evaluation systems and trust, Gregg and Scott (2006) emphasized the role of evaluation systems in reducing online auction fraud, Pavlou and Dimoka (2006) discussed the impact of text comments on trust building, price premiums, and seller differentiation, and Zhou & Windle (2008) designed a comprehensive framework for assessing the effects of feedback-based evaluation system in online marketplaces. These extant studies confirmed the relationship between evaluation systems and trust building. Regardless of their usefulness in building trust, there are intrinsic limitations of evaluation systems due to the difficulty in the presentation of accurate feedbacks (Resnick et al., 2000). In addition, the importance of buyers’ evaluations has prompted evaluation manipulations by some sellers. Sellers’ evaluation manipulation is not only observed on Taobao.com but also on well-known online marketplaces, such as eBay.com and Yahoo.com. Most online marketplaces cannot deter or disclose these manipulation behaviors effectively. The result of manipulations by sellers is marked by large amounts of positive feedback and few negative ratings, creating unreliable transaction environments and fragmented information which can influence consumers’ judgments (Resnick & Zeckhauser, 2002).

An online marketplace evaluation system is usually designed for users to express their experiences by numerical ratings and text comments. Higher rating scores and favorable comments appeal to buyers. Intuition suggests that rating scores are more prone to
manipulations than text comments because they are just numbers, while text comments need to be entered word by word. Meanwhile, rating scores can be added up conveniently and the average score presented by the system readily. On the contrary, it is hard to quantify the “average score” represented by text comments; the meaning of text comments need to be interpreted by users individually. Generally speaking, higher rating scores immediately attract buyers’ favorable attention because the information it carries is easily comprehensible. But Pavlou and Dimoka (2006) emphasized that buyers would intentionally read text comments, because text comments usually contain more information concerning the quality of products than numerical ratings. It is interesting to investigate whether buyers are getting accurate feedback information through the rating score versus the text comment mechanism, especially when they are under the influence of seller manipulations. In addition, novice or experienced online shoppers may not be equally susceptible to seller manipulations. Thus in constructing a model to investigate the effect of seller manipulation on buyer evaluations, we would consider buyers’ online shopping experience as a moderator of the relationship between manipulation and evaluation.

LITERATURE REVIEW

Online Marketplaces

Online marketplaces are generally categorized as B2B, B2C, and C2C models. For example, Alibaba.com is known for B2B e-commerce, Amazon.com for B2C e-commerce, and eBay.com and Taobao.com for C2C e-commerce. Although there are more well-known B2C e-commerce players, the development of C2C e-commerce has gradually grown to reach an unprecedented scale. With B2C and C2C players aggressively expanding their territories, the boundary of B2C and C2C is becoming blurry. B2C player like Amazon.com now embeds C2C model within its B2C platform, and C2C player like Taobao.com embeds B2C models within its auction platform. C2C online marketplace provides convenient transactional environments for individual sellers and buyers to exchange goods and share information. Individual sellers nowadays typically are small and medium businesses/enterprises (SMEs) rather than individual consumers who are also interested in offering items online.

SMEs can set up their store fronts easily in online marketplaces. The lower entrant cost attracts both SMEs and individual sellers. Brand equity is not what these small sellers possess. They are typically unknown and unrecognizable to potential buyers. The items that they sell on the online marketplace may not carry a known brand either. To harness sale opportunity, they need to build buyers’ trust through various trust building mechanisms such as third-party certification, evaluation system, and return policy. Studies showed that buyer’s positive feedback in online evaluation systems effectively attract other potential buyers’ attention (Chang, Cheung, & Tang, 2013). Payments are made by buyers without much assurance from sellers. The assurance is merely formed by buyers’ own assessment of sellers’ reputation and perhaps third-party payment mechanism which holds the payment as an escrow account until items are received (Lee & Lin, 2013). Therefore, online evaluation system plays an important role in buyers’ assessment of sellers’ reputation. Seller’s trustworthiness is judged by viewing buyers’ evaluations of their transaction experiences. Continuous transaction is sustained by the induced trust. Online evaluation systems provide reference information to help consumers make purchase decisions and promote the growth of C2C e-commerce (Ba & Pavlou, 2002). Therefore, the incentives for sellers to manipulate online evaluations in online marketplaces are the chances to become reputable sellers and to attract more buyers (MacInnes et al., 2005).
Evaluation Mechanism and Evaluation Systems

Trust-building mechanisms are indispensable on online marketplaces because online transactions lack face-to-face contact and that unknown smaller sellers make up the supply side. Information transparency and impartial transactions are not properly perceived if trust is absent. Li (2010) pointed out that seller’s credibility is indicated in evaluation systems, which result in more BIN transactions. Sellers are extremely apprehensive for their rating scores because ratings seriously impact the successful rate of transactions (Melnik & Alm, 2002; Song & Baker, 2007; Dini & Spagnolo, 2009).

Most B2C and C2C online marketplaces have implemented evaluation systems not only to help sellers build their credibility, but also to promote the brand image of the online marketplace. Resnick et al. (2000) defined online evaluation systems as online electronic systems that collect, distribute and aggregate feedbacks about participants’ past behavior. Buyers rely on evaluation systems to gauge sellers’ credibility; online marketplace operators rely on them to eliminate untrustworthy sellers. There are usually two parts of evaluations, including numerical ratings and text comments. In the text comments of evaluations, buyers can express their opinions about the transaction process and the quality of product/service received on the online evaluation system. The numerical ratings of evaluations are openly accumulated to represent the seller’s credibility and to be consulted when consumers make purchase decisions (Pavlou & Dimoka, 2006). For example, as shown in Figure 1, on Taobao.com, ratings are expressed as positive, neutral and negative. The accumulated rating score and text comments are also displayed.

Previous studies have confirmed the benefits of evaluation systems when they are properly designed and managed to provide reliability information for online consumers (Dellarocas, 2003). On the contrary, the viability of evaluation systems is challenged when the system face the problems of online transaction disputes, customer complaints and manipulation behavior (Gregg & Scott, 2008). Among them, manipulation behavior can be the most troublesome, because under manipulation, there is the concern that buyers’ opinions are not accurately expressed and other problems are essentially hidden.

Evaluation Manipulation

Online evaluations affect buyers’ choices of sellers, buyers’ final purchasing prices and the probability of successful transactions (Huang et al., 2011), and in turn impact sellers’ net revenue (Song and Baker, 2007). The importance of online evaluations has triggered various approaches of sellers’ manipulation to encourage buyers to leave positive ratings (Melnik & Alm, 2002), as shown in Figure 2. To avoid negative evaluations, some sellers disguise themselves as buyers to leave positive comments, while some urge buyers to alter their evaluations through using intimidating language (Hu et al., 2012; Hu, Liu, & Sambamurthy, 2011). Opportunistic sellers can be quite innovative in coming up with manipulation schemes (Dini & Spagnolo, 2009), including favorable rating persuasion, feedback theft, and reputation purchase (Dini & Spagnolo, 2009).

The most common practice of evaluation manipulation enacted by sellers is to constantly post reminders to attempt to influence buyers’ evaluation. In general, sellers’ reminder manipulation usually include “give us full stars and then you will get a discount,” “do not leave negative or neutral evaluations without trying to resolve the problem with us,” and “if you intend to leave a negative evaluation without contacting us firstly, please do not purchase anything here.” The spread of sellers’ reminders ranges from prompting a positive rating to prohibiting a negative rating. The type of sellers’ manipulation can be categorized into allurement and rejection. This study focuses on sellers’ evaluation manipulations that attempt to influence
buyers’ evaluations through posting reminders. Buyers’ comments and ratings can be restricted and not accurately expressed, if buyers are influenced or even intimidated by sellers’ reminders.

Figure 1: Example of seller’s evaluations on Taobao.com
Huang et al. Sellers’ Manipulation Influence Buyers’ Evaluations

Figure 2: Example of seller’s manipulation behavior on Taobao.com

(Note: The statement in red box is translated as follows:
"We refuse novice buyers who directly post negative ratings without fully understanding the product information or communicating with us beforehand. Please do not give negative or neutral ratings without communicating with us first. Although neutral ratings do not cause a reduction in rating score, they still reduce the percentage of positive ratings.

Percentage of Positive Ratings and Comment Orientation

In online marketplaces, sellers usually care more about numerical ratings than text comments, because numerical ratings are easily comprehended by buyers and higher rating scores imply credibility. Each positive rating contributes to an increment of accumulated rating score; whereas each negative rating leads to a decrement. Although neutral ratings have not significantly influenced scores, they dilute the percentage of positive ratings. On Taobao.com, sellers’ credibility is further ranked by the e-marketplace operator using the accumulated rating scores, with different ranks marked by different symbols like crowns, diamonds, etc. A higher rank implies a higher level of credibility, setting a clear direction for sellers to strive to earn more positive ratings (and as few neutral and negative ratings as possible) in order to advance in credibility level. However, the accumulated numerical rating is directly associated with how long a seller has been conducting business in the marketplace; it does not accurately reflect a seller’s current situation. The percentage of positive ratings reflects a seller’s recent dealings. It is calculated by taking recent data, say six months in Taobao.com’s case, and dividing the number of positive ratings by the total number of ratings. To observe whether sellers’ manipulation affects buyers’ evaluations, this study chooses the percentage of positive ratings, rather than the accumulated rating score, as the measurement of numerical rating scores, because it is an indicator of more recent ratings.

Besides numerical rating scores, text comments constitute the other important part of rating. Although the information they present is not as direct or obvious as numerical ratings, text comments contain much realistic information that cannot be properly reflected in numerical ratings. Pavlou and Dimoka (2006) pointed out that browsing rich text comments not only shapes buyers’ trust in sellers’ trustworthiness and benevolence, but also boosts the goodwill of online marketplaces. According to Pavlou and Dimoka (2006), about 97% of consumers read
buyers’ comments and 81% read at least 25 comments (Pavlou and Dimoka, 2006). When giving an evaluation of a seller, a positive numerical rating can avoid offending the seller directly. But the buyer can always turn around and write negative comments to express their dissatisfaction. Inconsistent evaluations between text comments and rating scores do exist. The inconsistency is not easily detected by sellers, and they challenge buyers’ discernment. Besides numerical rating scores, comment orientations also influence buyers’ purchase decision (Kusumasondjaja et al., 2012; Park & Lee, 2009). It is of great interest to understand whether they are affected under sellers’ manipulation.

**Buyer Experience**

Gregg and Scott (2006) observed that experienced buyers are more capable of effectively utilizing evaluation systems to analyze sellers’ reputation and avoid online fraud. Also, experienced buyers realize the importance of evaluations and tend to be more cautious about the evaluations they provide. They are more familiar with the way evaluation systems work, therefore could be shrewd in shunning disputes with sellers. Evaluation systems usually allow two-way evaluations, thus there is a clear benefit for buyers to keep peace with sellers in exchange of sellers’ favorable evaluations on them as buyers (MacInnes et al., 2005). The overall atmosphere of an e-marketplace can partially determine how much truth buyers are willing to reveal.

Buyers’ online behavior is closely associated with their online experience. It is reasonable to speculate that whether buyers express their true opinions on evaluation systems is also linked to their experience shopping on the e-marketplace. Therefore this study adopts buyer experience as the moderator in studying the effect of seller manipulation. Due to the difficulty of reaching each seller’s buyers through the buyer itself or the e-marketplace operator, this study use the accumulated rating (as rated by all sellers the buyer had completed transactions with) of each buyer as a surrogate for buyer experience.

**RESEARCH METHOD**

The purpose of this study examines the influence of sellers’ manipulation on buyers’ numerical ratings and text comments in online evaluation systems while buyers’ experiences is a moderator. This study also investigates the consistency between the sellers’ percentage of positive ratings and orientation of text comments. In this section, the hypotheses and research model are described in detail and introduce the process of data collection, sampling, measurement, and data analysis respectively.

**Hypotheses and Research Model**

This study considers the type of systematic seller manipulations of posting reminders to ask buyer to give positive rating or not to leave negative/neutral rating. Manipulations are classified into two categories: allurement and rejection. Allurement manipulation means that sellers post reminders to offer discounts if buyers leave positive ratings. Rejection manipulation means that sellers seek to reject transactions if buyers do not comply with sellers’ expectation to leave positive feedback. Therefore the possible values of the variable, manipulation, are rejection, allurement, and none. The percentage of positive ratings and the orientation of text comments are the dependent variables in interest. The research model is shown in Figure 3.
The following null hypotheses are tested.

H1: There is no significant relationship between seller manipulation and the percentage of positive ratings.

H2: There is no significant relationship between seller manipulation and the orientation of buyers’ comments.

H3: Buyer experience does not moderate the relationship between buyer evaluation and seller manipulation.

H3a: Buyer experience does not moderate the relationship between seller manipulation and the percentage of positive ratings.

H3b: Buyer experience does not moderate the relationship between seller manipulation and the orientation of buyers’ comments.

**Sampling and Data Collection**

Empirical data collected from Taobao.com is used to test the proposed research model. Keyword search is used as a filtering mechanism to select sellers who have been continuously doing business on Taobao.com for at least six months from the time data is collected. For Taobao.com is a Chinese C2C marketplace, a Chinese keyword table that is drawn from a Chinese character frequency survey (Ho, 1998) is referenced to choose suitable keywords. From the website, a list of 210 sellers is compiled; they are classified into three groups: allurement, rejection, and none, based on the content that sellers put on the product pages. The data collected from the virtual store of each seller include the manipulation type (allurement, rejection, none), percentage of positive ratings, and the most recent twenty-five text comments left by buyers (Pavlou and Dimoka, 2006). Buyers who left these twenty-five comments are also identified, and the buyer experience is estimated based on their average accumulated rating as buyers. Based on the median of buyer experience, two groups with high and low experience are formed.

5,250 buyers’ comments are collected (25 comments times 210 sellers). The orientation of text comments for each seller is assessed by coders recruited from online communities and bulletin board systems. Each set of 25 comments is coded by 5 coders; the orientation of these 25 text comments is assessed by averaging coders’ responses to an online web questionnaire which measures text comment orientation. Coders are randomly sampled online and screened.
to assure that they are experienced online shoppers and are familiar with the operation of online marketplaces. There are 1050 coders (5 coders times 210 sellers).

Measurement

As described previously, all data except comment orientation are extracted from the website of Taobao.com. The measurement scale of comment orientation is constructed according to the dimensions of service quality, information quality, product quality, and buyers’ satisfaction. For service quality, 7 items of Chong (2004) are adopted. Purohit and Srivastava (2001) and Zaichkowsky (1994) are the sources for the scales of information quality (4 items), buyer’s satisfaction (4 items), and product quality (5 items).

Pretesting and Items Modification

Thirty-four coders participated in the pretest of the measurement scale. SamrtPLS2.0 (Ringle, Wende, and Will, 2005) is used to analyze the pretest data. Items are removed according to factor loading, average variance extracted (AVE), and composite reliability (CR). The suggested thresholds by Hair et al. (1998) Bagozzi & Yi (1988), and Fornell & Larcker (1981) are used, which 0.7, 0.5, and 0.7 respectively. The reliability is tested by Cronbach’s α and the threshold is 0.7, as suggested by Cortina (1993).

After aforementioned analysis, the convergent validity of four constructs has been satisfied to achieve all thresholds and nine items are removed from questionnaire including four items in service quality, one item in information quality, two items in product quality and buyer’s satisfaction. The discriminant validity is also tested and the condition that the square root of AVE should be larger than the corresponding inter-construct correlations (Fornell and Larcker’s 1981) is met.

ANALYSIS AND RESULTS

SmartPLS software is used to conduct measurement model analysis to examine the reliability and convergent and discriminant validity. Multivariate analysis of variance (MANOVA) is used to test hypotheses, and the consistency between the percentage of positive ratings and the comment orientation is examined by Pearson correlation analysis. Table 1 shows the result of reliability and validity analysis. All conditions of high reliability and validity are met, except a factor loading (0.68) of one item, which is slightly below the threshold of 0.7. Table 2 shows the result for discriminate validity analysis; all constructs have satisfied the condition that the square root of AVE is larger than the corresponding inter-construct correlations.

Hypothesis Testing

Pearson correlation analysis between the percentage of positive ratings and comment orientation shows that under the influence of seller manipulation, the consistency between them is quite low, as the correlation coefficient is only 0.169 (p<0.01). The low correlation coefficient implies that buyers may give a positive rating but not necessarily write a positive comment.

After Pearson correlation analysis, this study intended to use multivariate analysis of variance (MANOVA) to analyze the effect of sellers’ manipulation behavior on buyers’ evaluations, because it can compare several criteria variables simultaneously among groups. According to Hair et al. (1998), MANOVA is especially sensitive to outliers which have a disproportionate effect on the analysis results. Therefore, this study examines the outliers and then eliminates participants of 6 sellers and 6 coders. Meanwhile, the Levene’s test of
homogeneity in the percentage of positive ratings is significant (F-value=30.099, p<.001). This violates the assumption of MANOVA and represents sampling heterogeneity, thus MANOVA test is replaced with analysis of variance (ANOVA) test.

Two-way ANOVA is used to analyze the relationship between continuous dependent variables and two categorical independent variables. The two categorical independent variables are seller manipulation behavior and buyer experience. Although the homogeneity (F-value=30.099, p<.001) assumption of ANOVA is not met, Hair et al. (1998, p.347) indicates that F tests in ANOVA are robust with regard to the assumptions except in extreme cases.

Table 1: Measurement model

<table>
<thead>
<tr>
<th>Construct/ Item</th>
<th>Mean</th>
<th>Standard error</th>
<th>Factor loading</th>
<th>AVE</th>
<th>Composite reliability</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service quality (SQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ 1</td>
<td>5.00</td>
<td>1.214</td>
<td>0.911</td>
<td>0.847</td>
<td>0.943</td>
<td>0.914</td>
</tr>
<tr>
<td>SQ 2</td>
<td>5.00</td>
<td>1.203</td>
<td>0.902</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ 3</td>
<td>4.91</td>
<td>1.233</td>
<td>0.948</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information quality (IQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ 1</td>
<td>4.82</td>
<td>1.299</td>
<td>0.868</td>
<td>0.778</td>
<td>0.913</td>
<td>0.861</td>
</tr>
<tr>
<td>IQ 2</td>
<td>4.63</td>
<td>1.234</td>
<td>0.916</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ 3</td>
<td>5.09</td>
<td>1.119</td>
<td>0.861</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product quality (PQ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PQ 1</td>
<td>5.09</td>
<td>1.287</td>
<td>0.803</td>
<td>0.779</td>
<td>0.913</td>
<td>0.909</td>
</tr>
<tr>
<td>PQ 2</td>
<td>5.05</td>
<td>1.275</td>
<td>0.873</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PQ 3</td>
<td>4.56</td>
<td>1.241</td>
<td>0.965</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buyer satisfaction (SAT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAT 1</td>
<td>5.48</td>
<td>1.244</td>
<td>0.978</td>
<td>0.709</td>
<td>0.825</td>
<td>0.900</td>
</tr>
<tr>
<td>SAT 2</td>
<td>5.52</td>
<td>1.207</td>
<td>0.680</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The results for discriminate validity

<table>
<thead>
<tr>
<th>Construct</th>
<th>Service quality</th>
<th>Information quality</th>
<th>Product quality</th>
<th>Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service quality</td>
<td>0.921</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information quality</td>
<td>0.795</td>
<td>0.882</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product quality</td>
<td>0.669</td>
<td>0.762</td>
<td>0.883</td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td>0.558</td>
<td>0.565</td>
<td>0.603</td>
<td>0.842</td>
</tr>
</tbody>
</table>

Percentage of Positive Ratings

The effect of seller manipulation on the percentage of positive ratings, as moderated by buyer experience, is examined through two-way ANOVA. The analysis confirms the main effect, i.e., seller manipulation significantly influences the percentage of positive ratings (F-value=53.839, p<.001) and therefore H1 is rejected. The interaction of seller manipulation and buyer experience also significantly influences the percentage of positive ratings (F-value=13.641, p<.001) and therefore H3a is also rejected, as shown in Table 3. Two-factor interaction effect of sellers’ manipulation and buyers experience on the percentage of positive ratings is shown in Figure 4. For the group of experienced buyers, regardless of sellers’ manipulation, sellers enjoy higher percentage of positive ratings.

As can be interpreted from Figure 4 and confirmed by the Post Hoc tests shown in Table 4, for the group of inexperienced buyers, different types of manipulation lead to different percentage of positive ratings. Apparently inexperienced buyers respond negatively to manipulations, because sellers receive higher percentage of positive ratings when they do not
intend to manipulate ratings. The allurement approach by sellers results in the lowest percentage of positive ratings.

Table 3: Two-way ANOVA test of percentage of positive ratings

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>168.940</td>
<td>5</td>
<td>33.788</td>
<td>58.145</td>
<td>.000</td>
</tr>
<tr>
<td>Seller manipulation</td>
<td>62.571</td>
<td>2</td>
<td>31.286</td>
<td>53.839</td>
<td>.000</td>
</tr>
<tr>
<td>Buyer experience</td>
<td>60.620</td>
<td>1</td>
<td>60.620</td>
<td>104.320</td>
<td>.000</td>
</tr>
<tr>
<td>Buyer experience x Seller</td>
<td>15.853</td>
<td>2</td>
<td>7.927</td>
<td>13.641</td>
<td>.000</td>
</tr>
<tr>
<td>manipulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: The effect of the interaction on the percentage of positive ratings

![Graph showing the effect of the interaction on the percentage of positive ratings]

Table 4: Scheffé’s multiple comparisons of percentage of positive ratings

<table>
<thead>
<tr>
<th>Seller manipulation (I)</th>
<th>Seller manipulation (J)</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allurement</td>
<td>None</td>
<td>- .7098*</td>
<td>.05768</td>
<td>.000</td>
<td>-.8512 - .5685</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rejection</td>
<td>-.4746*</td>
<td>.05743</td>
<td>.000</td>
<td>-.6154 - -.3338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Allurement</td>
<td>.7098*</td>
<td>.05768</td>
<td>.000</td>
<td>.5685 - .8512</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rejection</td>
<td>.2353*</td>
<td>.05813</td>
<td>.000</td>
<td>.0928 - .3777</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rejection</td>
<td>Allurement</td>
<td>.4746*</td>
<td>.05743</td>
<td>.000</td>
<td>.3338 - .6154</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>-.2353*</td>
<td>.05813</td>
<td>.000</td>
<td>-.3777 - -.0928</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

Comment Orientation

Two-way ANOVA is also used to examine the effect of seller manipulation on comment orientation, as moderated by buyer experience. Levene’s test of homogeneity is insignificant (F-value=1.010, p>.05), which meets the assumption of ANOVA. As shown in Table 5, the main effect is not confirmed, i.e., sellers’ manipulation does not influence the comment orientation (F-value=1.235, p>.05), and therefore H2 is accepted. But the interaction of seller manipulation and buyer experience significantly influences the comment orientation (F-value=3.669, p<.05) and therefore H3b is rejected.
Two-factor interaction effect of seller manipulation and buyer experience on the comment orientation is shown in Figure 5. Again, the result clearly shows that if sellers take the rejection approach to manipulate buyers’ evaluation, they will get the most undesirable evaluation, as the percentage of positive rating is lowest regardless of buyer experience. The allurement approach apparently invite more positive ratings from experienced buyers, but more negative ratings from inexperienced buyers. The clear crossing of the two lines in Figure 5 demonstrates the strong and opposite effects. The hypothesis testing results are shown ummed upvalues and significance levels marked.re 6, with thee strong and opposite effects. more negative ratings from in experieshownsh in Figure 6, with F-values and significance levels noted.

Table 5: Two-way ANOVA test of comment orientation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>d.f.</th>
<th>MS</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>9.484</td>
<td>5</td>
<td>1.897</td>
<td>1.827</td>
<td>.105</td>
</tr>
<tr>
<td>Seller manipulation</td>
<td>2.565</td>
<td>2</td>
<td>1.283</td>
<td>1.235</td>
<td>.291</td>
</tr>
<tr>
<td>Buyer experience</td>
<td>.063</td>
<td>1</td>
<td>.063</td>
<td>.061</td>
<td>.805</td>
</tr>
<tr>
<td>Buyer experience x Seller manipulation</td>
<td>7.619</td>
<td>2</td>
<td>3.809</td>
<td>3.669</td>
<td>.026</td>
</tr>
</tbody>
</table>

Figure 5: The effect of the interaction on comment orientation
DISCUSSION

For understanding the relationship between the percentage of positive ratings and the comment orientation, the correlation analysis is adopted and showed a low correlation between them. The result confirms that evaluating numerical ratings and text comments by buyers are isolated. Moreover, according to aforementioned hypothesis tests, sellers’ manipulation behavior directly influences buyers’ evaluations, especially impacts on the percentage of positive ratings that is a critical reference for buyers to make purchase decision. Many of consumers make a deal with a seller based on seller’s rating scores (Huang et al., 2011; Melnik & Alm, 2002), because the numerical ratings are playing an important with explicit direction for potential buyers to obtain the most easily understandable evidence of sellers’ credibility (Ba & Pavlou, 2002; McDonald & Slawson, 2002). Therefore, this study speculates that not only sellers’ credibility in the buyers mind is based on numerical ratings, but also improving sellers’ credibility to enhance exchange opportunities is the main reason why sellers try to manipulate evaluations. However, text comments are not affected by sellers’ manipulation behavior to truthfully describe buyers’ perceptions of transaction process and contain sincere with plentiful information for prospective buyers (Pavlou & Dimoka, 2006). That is, buyers’ comment orientation is not influenced by sellers’ manipulation behavior.

In addition, the interaction between sellers’ manipulation behavior and buyers’ experience can affect buyers’ evaluation including percentage of positive ratings and comment orientation. Regarding the comment orientation, the more experienced buyers tend to write more positive evaluation than the less ones in allurement and rejection group. Especially, the higher experienced buyers in allurement group are more obvious than the ones in rejection group. Meanwhile, the higher experienced buyers in allurement group tend to accept sellers’ manipulation conditions including discount or reward (Tong et al., 2013). However, the less experienced buyers who are more easily satisfied give more positive comments than the higher ones who evaluate transaction process seriously in the none manipulation group.

CONCLUSIONS

This study examines the effect of seller manipulations on buyers’ online evaluations. The relationship is investigated by considering buyer experience as a moderator. Seller manipulation
is operated as having three levels: none, allurement, rejection. Two aspects of evaluations are considered: rating score and text comments. By using empirical data from the website of Taobao.com, the analysis results show that manipulations do not bring sellers any benefits in terms of more favorable ratings from inexperienced buyers. However, and unfortunately, experienced buyers do tend to give more favorable comments under sellers’ allurement type of manipulations. Also, in general, experienced buyers are more likely to give positive rating scores than inexperienced buyers. In the case of inexperienced buyers, seller manipulations hurt both the percentage of positive ratings and the comment orientation. With respect to the percentage of positive ratings, inexperienced buyers respond to the rejection type of manipulations more mildly than the allurement type of manipulations. Allurement type of manipulations brings sellers lowest rating scores in terms of the percentage of positive ratings. It is puzzling that rejection type of manipulations do not hurt as much, compared with allurement type. Perhaps rejection type of manipulation is intimidating to these buyers, so sometimes they were hesitant to give low rating scores.

The correlation coefficient between percentage of positive ratings and comment orientation is low, implying that buyers act very differently when doing these two types of evaluations. This indirectly explains the dissimilar effects manipulations have on rating scores and comment orientation. Although this study has interesting findings, it has also a few limitations that should be noted. First, the data is obtained from only one e-marketplace, Taobao.com. Although the scale of Taobao.com mandates its worthiness of study, future research should confirm the similar results from more e-marketplaces and perhaps other cultural regions. Second, there are many faces of seller manipulations, but this study considers only the typical systematic manipulation of seller posting allurement or rejection messages. Future research can probe deeper into this issue by adopting case study research method to enrich the findings. Third, this study does not try to solve the problem of seller manipulations; instead it merely unveils effects of seller manipulations. With the interesting results brought forth by this study, we hope to trigger the interest of e-marketplace operators to strive for better, fairer, and more transparent online evaluation systems.

REFERENCES

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and
measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 328-388.


