

RESEARCH

Open Access

How buyers perceive the credibility of advisors in online marketplace: review balance, review count and misattribution

Kewen Wu^{*}, Zeinab Noorian, Julita Vassileva and Ifeoma Adaji

^{*} Correspondence:
kew259@mail.usask.ca
Department of Computer Science,
University of Saskatchewan,
Saskatoon, Canada

Abstract

In an online marketplace, buyers rely heavily on reviews posted by previous buyers (referred to as advisors). The advisor's credibility determines the persuasiveness of reviews. Much work has addressed the evaluation of advisors' credibility based on their static profile information, but little attention has been paid to the effect of the information about the history of advisors' reviews. We conducted three sub-studies to evaluate how the advisors' review balance (proportion of positive reviews) affects the buyer's judgement of advisor's credibility (e.g., trustworthiness, expertise). The result of study 1 shows that advisors with mixed positive and negative reviews are perceived to be more trustworthy, and those with extremely positive or negative review balance are perceived to be less trustworthy. Moreover, the perceived expertise of the advisor increases as the review balance turns from positive to negative; yet buyers perceive advisors with extremely negative review balance as low in expertise. Study 2 finds that buyers might be more inclined to misattribute low trustworthiness to low expertise when they are processing high number of reviews. Finally, study 3 explains the misattribution phenomenon and suggests that perceived expertise has close relationship with affective trust. Both theoretical and practical implications are discussed.

Keywords: Source credibility; Misattribution; Online marketplace; Review balance

Introduction

In an online marketplace, buyers rely heavily on reviews posted by advisors. A recent business survey reported that 92% of online consumers read advisors' reviews before they make purchase decisions [1]. Literature also suggests that advisors' reviews significantly influence consumers' attitudes towards the products or sellers, which ultimately influence sales [2,3].

The extent to which a buyer accepts or follows an opinion presented in a review is a matter of persuasiveness. The persuasiveness of an online review is determined by the credibility of its source (the advisor), because online reviews are written by advisors with varied backgrounds and motivations [4]. Advisors can write reviews no matter if they are capable of assessing a product critically or not (e.g., layperson versus expert). Moreover, many intentional and unintentional factors can influence the writing of a review [5-7]. For instance, an advisor's account may be controlled by a seller to write

positive reviews and promote himself (known as ballot stuffing); and it may also be controlled to write negative reviews to attack competitors (known as bad-mouthing). These reputation manipulation activities have been identified as a pervasive phenomenon in online marketplaces [5,8]. Even if an advisor is a real buyer, he may still be influenced by others and write reviews that do not represent his actual experience (e.g., herd effect).

Given the uncertainty regarding the source of online reviews, buyers are motivated to assess the credibility of advisors based on accessible pieces of information [9]. Many online marketplaces (e.g., Amazon, Taobao) allow buyers to visit advisors' profile page. To evaluate an advisor's credibility, buyers are inclined to seek and use profile information as cues, other than the review itself. A number of studies have been conducted to evaluate how advisors' profile influences buyers' perception of credibility [10,11]. Advisors' static profile information, such as real name, location, nickname and hobbies, have been found to be helpful in supporting consumers' judgment [11,12]. However, current studies on advisors' review history mainly come from computer science field, and little is known about the impact of advisors' review history on buyers' perception of advisors' credibility. Analyzing an advisor's review history could provide useful information (e.g., purchase frequency, areas of interests or even background) about the advisor, which can be helpful for buyers to make judgement on advisors' credibility.

In this paper, we segment advisors into five types based on the ratio of positive to negative reviews (referred to as review balance). If the proportion of positive (negative) reviews is extremely higher than, substantially higher than, or almost equal to the proportion of negative (positive) reviews, the review balance is respectively defined as extreme positive (negative), positive (negative), or neutral. We choose review balance as representative of review history because it can be easily noticed by buyers through direct scanning of an advisor's review history list or a summary table provided by the platform. Prior studies indicate that buyers usually do not scrutinize reviews [13,14]; they form attitude only based on the information they gain easily. Intuition also suggests that it is unrealistic for a buyer to conduct a comprehensive evaluation of review history for each advisor in the product page.

We conducted three sub-studies to explore how different review balances signal different meanings to buyers regarding the advisors' trustworthiness and expertise (two dimensions of credibility). Study 1 aims to gain a preliminary knowledge about buyers' perception of advisor's trustworthiness and expertise. Study 2 extends study 1 by using larger sample size and considering more variables. Finally, study 3 is conducted to further explain the results of previous two sub-studies.

Research background

Source credibility: trustworthiness and expertise

The concept of source credibility has received much attention from various fields related to communication, such as politics, human-computer interaction, marketing and information system. It is a multifaceted term suggesting that the positive characteristics of a message source can enhance the perceived value of message information, and thus increase the persuasiveness of the message [15,16]. Expertise, trustworthiness and attractiveness are commonly reported as three dimensions of source credibility [17]. In this study, we considered source credibility as a two-dimensional construct, since expertise and trustworthiness

are more relevant to online review context [18]. Trustworthiness describes the receiver's confidence in a source's objectivity and honesty in providing information [15]. There is a wide consensus on the positive relationship between trustworthiness and source credibility [19].

Expertise refers to a source's capability of providing correct and valid information [15]. Such capability can be technical-oriented or practical-oriented [20]. Technical expertise reflects the skillfulness of processing special knowledge required by writing comments towards a given product (e.g., an advisor who majors in acoustics writes a review about a headphone). Practical expertise is the skills that are gained from direct participation in related activities (e.g., an advisor who has tried many headphones writes a review about one headphone). The characteristics of online communication (e.g., limited availability of personal information) make it difficult to identify whether an advisor is an expert or not. As a result, in online context, different results have been found regarding the relationship between expertise and source credibility. For example, some studies found that expert endorsers can lead to higher source credibility than laypersons; others found that layperson can induce higher credibility than experts; yet others found that the levels of expertise make no difference in determining the perceived source credibility [19,21].

The complex findings on expertise imply that other dimensions of source credibility might disturb the effects of expertise. As mentioned earlier, attractiveness is not relevant to online review context. Here we only take trustworthiness as an example. On one hand, high expertise can lead to increased trust because assessments of expertise and trust both employ an attribute evaluation of trustee's identifiable actions [22]. For example, a seller's expertise reflects a buyer's identification of competencies associated with the transaction. On the other hand, as suggested by the attribution theory [23], people attribute a review to both stimulus and non-stimulus causes. When the consumer suspects that the review is not drawn based on product performance (stimulus) but on the advisors' unknown intentions (non-stimulus), they will discredit the review message. In some cases, a source may be perceived to be high in expertise but low in trustworthiness [24]. For example, people trust an expert because they think expert statements are true; however, if this expert's motivation to share is reasonably suspected, people's perception of this expert's trustworthiness will decrease. The contradictory effects (e.g., high on expertise but low on trustworthiness) may cancel each other out [25].

The above mentioned two circumstances only address the impacts of expertise on trustworthiness, that high expertise can lead to both high trust (because of belief in competency) and low trust (because of suspicious motivation). However, little is known about how trustworthiness affects expertise.

Advisors profile and credibility

Previous work on credibility of online reviews can be divided into two streams. The first stream of work focuses on review itself; studies have addressed many factors such as sequence of reviews [26,27], valence [26], volume [28], information depth [29], attribution (e.g., experience issue or product issue) [26,27]. However, these studies generally assume reviews come from credible sources.

The second stream of work deals with the credibility of advisors. Much work has been done on evaluating the effects of advisors' profile. In real online review systems, a profile usually includes an advisor's identity-related information and review history. Advisors' identity-related information, such as real name, gender, location, nickname, hobbies and reputation (e.g., special badges such as top 50 reviewers), has been proven to be helpful for buyers' judgment [11,12,10]. However, limited attention has been paid on the effects of review history.

The social exchange theory suggests that people develop trust based on behavioral characteristics observed from direct experiences with the trustee [30]. The history of experience facilitates the accumulation of knowledge and thus increases the validity of knowledge-based attribution [31]. Compared to static characteristics (e.g., gender, location), buyers are able to make rational credibility judgment as they obtain greater knowledge from the review history.

Positive or negative reviews could signal different meanings to buyers, for instance, a reviewer who gives negative feedback might be perceived to be high in expertise [32]. However, few studies have considered how buyers perceive expertise from advisor's review history (e.g., review balance). Moreover, current studies on the perception of trustworthiness from advisors' review history mainly come from computer science area. The basic assumptions regarding trustworthiness and advisors' review behavior are based on three points: (1) Similarity. According to social identity theory [33], a buyer may categorize an advisor who has similar purchase history and review opinions into the same social group, resulting in increased trust towards this advisor [34,35]. (2) Social consensus, that if an advisor holds the same opinions with the majority of advisors, his/her review is perceived as correct and would be accepted [36]. (3) Social network, that dishonest advisors (e.g., fake buyers' accounts), may share the same review behavioral pattern [37]. Given the fact that related human studies are scarce, this paper evaluates buyers' perception of advisors' credibility based on review history.

Data source

The review dataset used in this paper is built upon Taobao review data. We selected Taobao as our target online marketplace based on two reasons. First, Chinese online marketplaces have been growing rapidly in recent years. Taobao is the leading platform with about 90% market share. Its transaction volume is estimated to have more sales than Amazon and eBay combined in 2013 [38]. Taobao is well known among Chinese communities (half a billion registered users) and it is usually considered as a typical e-commerce sample in previous studies [39]. Second, despite the huge number of transactions, Chinese online marketplaces face serious reputation manipulation problem [5]. For example, some critics estimate that about 80% of *Taobao* sellers have committed reputation manipulation activities during their businesses [40]. And it has been reported that over 1000 active trust fraud companies provide services to help sellers increase reputation and whitewash negative feedback [5]. But a recent official report shows that more than 70% online buyers choose Taobao as their primary choice [41]. Therefore, the high transaction volume, serious trust issue and being buyers' primary choice jointly make Taobao a valuable target to investigate.

We use a self-developed crawler to download real review data from Taobao during 2014-04-01 and 2014-4-20. This dataset includes the latest 180-day detailed review information about 24,287 sellers and 1,686,870 advisors who are willing to show their profile. The average number of reviews per advisor in our dataset is 116.

To prepare the dataset for our experiment, we invited four master's students to select 200 positive and 200 negative reviews from our Taobao review database. The selection of reviews was based on two criteria: (1) previous studies have shown that the different review targets (product and service) have different impacts on consumer's decision-making process [26]. Therefore, we decided to only consider product attribute-based reviews to serve as data source in our experiment. Service-based reviews were excluded because service quality is usually unstable across different buyers (e.g., delivery service might be excellent in some areas but much worse in other areas) and buyers' perception of service quality contains many subjective factors. (2) We set the length of each review to be around 30 Chinese characters (about 60 English characters), and the reasons described in each review should be clear. We built advisors' profiles based on five types of review balances (See Table 1). In the following experiment, we did not set the ratio between number of positive ratings (R) and number of negative ratings (S) close to threshold values (e.g., 0.2 for Type I), because we wanted to make different types of review balance distinguishable. For example, we set the ratio of a Type I advisor's R/S to 0.05, rather than 0.19.

Study 1

Study 1 was designed to gain a preliminary knowledge about buyers' perception of advisor's source credibility regarding different review balances.

Hypotheses

Previous studies suggest that the proportion of positive reviews is much higher than negative reviews in online review systems [42,43]. People are reluctant to give negative feedback unless they encounter terrible experience [44]. A content analysis of eBay comments shows that 72.5% of negative reviews were related to unsatisfactory product and service, while the other 27.5% were related to sellers' attempts to exploit buyers [43]. This result suggests that terrible experience (negative feedback) usually happens due to the poor product or service quality that cannot meet buyer's expectation.

Table 1 Five types of advisors based on different review balance

Type	Description
I. Extremely negative Balanced	$R < <S^a$: number of positive ratings are significantly lower than number of negative ratings ($R/S < 0.2^b$)
II. Negative balanced	$R < S$: number of positive ratings are lower than number of negative ratings ($0.2 \leq R/S < 0.7$)
III. Neutral balanced	$R \approx S$: number of positive ratings are approximately the same as number of negative ratings ($0.7 \leq (R/S \text{ or } S/R) \leq 1$)
IV. Positive balanced	$R > S$: number of positive ratings are larger than number of negative ratings ($0.2 \leq S/R < 0.7$)
V. Extremely positive Balanced	$R > >S$: number of positive ratings are significantly larger than number of negative ratings ($S/R < 0.2$)

Note: ^a: R refers to number of positive ratings/reviews; S refers to number of negative ratings/reviews; ^b: this ratio is only used to describe a phenomenon (e.g., $R < <S$) and used to manipulate of advisors' profiles. It is not a strict classification of advisors.

The reviewers who give negative feedback are perceived as brighter and more intelligent than those who give positive feedback [32]. They give negative reviews because they have enough knowledge to identify product issues. For instance, as a domain expert, an acoustics enthusiast gives negative feedback to a headphone due to its poor performance, while non-experts could not notice the pros and cons of this headphone. In this view, an advisor with a negative review balance might be perceived as a strict expert who is hard to be satisfied. Therefore, we hypothesize that:

H1: The level of perceived expertise of an advisor increases as the review balance changes from extremely positive to extremely negative.

Negative feedback usually contains distinctive information than positive ones, therefore, it is perceived to be more accurate, trustworthy and helpful for buyers to make decisions [42]. Absence of negative feedback may have nothing to do with the judgment of review authenticity [19]. An advisor who has almost all positive feedback (review balance: extreme positive) may be considered as a malicious account controlled by a dishonest seller to do self-promotion, or as a “Mr. Goody-goody” who always gives positive feedback regardless of his actual experience. Similarly, an advisor who gives all negative feedback (review balance: extreme negative) may be judged to be a malicious account used to attack competitors, since the case that a buyer always experiences unsatisfactory transactions is unrealistic. Previous studies have found that buyers are more likely to form positive attitudes (e.g., trust, purchase intention) towards a product which receives a mix of positive and negative reviews [45,46,19]. Therefore, it is reasonable to assume that an advisor who posts both positive reviews and negative reviews would be perceived as trustworthy. We hypothesize that:

H2: The level of perceived trustworthiness is high when an advisor’s review balance is neutral, and the level of perceived trustworthiness is low when an advisor’s review balance is either extremely positive or extremely negative. Especially, an advisor with extreme negative review balance is perceived to be most untrustworthy.

Experiment and result

In order to reduce cognitive load, we only considered ratings in this sub-study. We created two sets of advisors’ profiles based on our review dataset. Advisors in each set have entirely different review balances (see Table 2). Although these advisors’ profiles cannot present the characteristics of the whole dataset, using a small amount of typical experiment material is acceptable in many studies [9,47].

Twenty experienced online buyers were invited to evaluate the impacts of review balance on perceived trustworthiness and expertise. These participants were all aware of

Table 2 Advisors’ profile used in study 1

Type	Description	Set 1 (R,S)	Set 2 (R,S)
I	$R < <S^a$	(5, 86), (0, 103)	(2, 42), (0, 63)
II	$R < S$	(31, 57), (38, 64)	(13, 30)
III	$R \approx S$	(51, 43), (58, 42)	(29, 24), (43, 32)
IV	$R > S$	(68, 31), (72, 23)	(37, 13), (64, 14), (56, 16)
V	$R > >S$	(104, 0), (115, 1)	(49, 1), (43, 1)

Note^a: R refers to number of positive ratings/reviews; S refers to number of negative ratings/reviews.

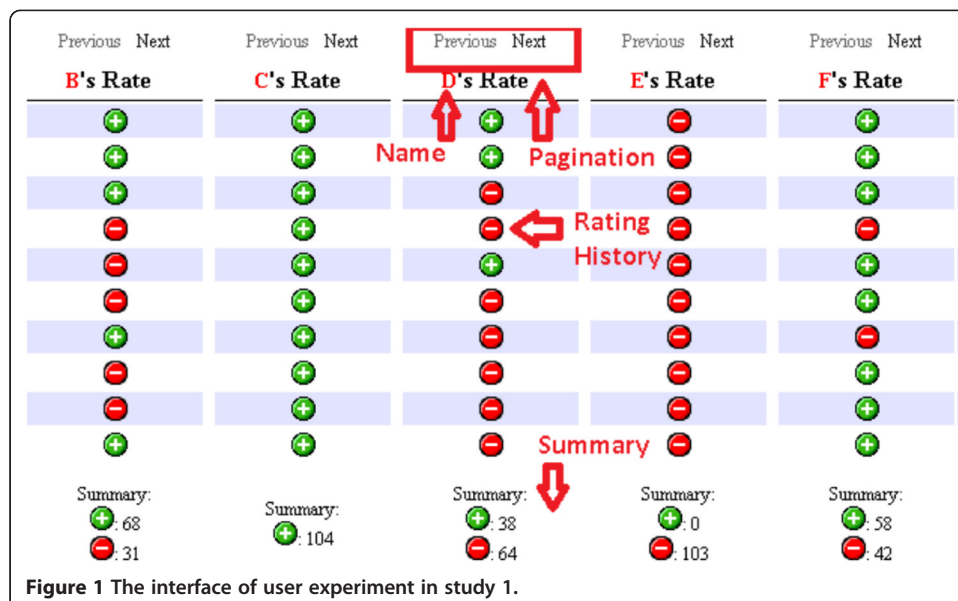
unfair rating/review phenomenon in online marketplaces, they were told that the rating history of each advisor in this survey was based on real data gained from Taobao. The interface of the experiment system is shown in Figure 1.

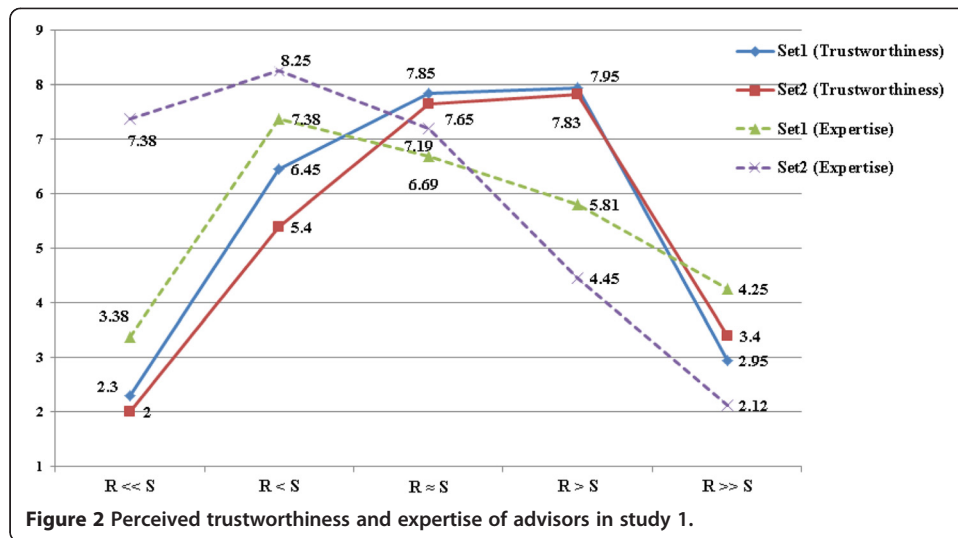
For the judgement of perceived trustworthiness, we randomly assigned 10 participants to check the rating history of advisors in Set 1 and asked them to rank advisors based on their perceived trustworthiness from the lowest (1) to the highest (10) on a ten-point scale (we used a computer program to ensure that each ranking position has only one advisor). Then we assigned the remaining 10 participants to rate advisors in Set 2 and rank advisors in the same way.

For the judgement of perceived expertise, we used the same advisors' profiles and the same subjects (however, two of them quitted). We randomly assigned 9 participants to check advisors in Set 1 and asked them to rank advisors based on perceived expertise from the highest to the lowest on the ten-point scale (1 shows the least expertise and 10 shows the highest expertise). Then we assigned the remaining 9 participants to check Set 2 and rank advisors, respectively.

We used Kendall's coefficient of concordance (W) to measure the degree of agreement among participants with the rankings of advisors. The capability of W in performing multiple judgments (more than two) makes it the most suitable tools to test inter-judge reliability [48]. Past studies suggest that the value of $W > 0.7$ shows strong consensus; $W = 0.5$ shows moderate consensus; and $W < 0.3$ shows weak consensus amongst different users on their ranked data [48].

In the test regarding perceived trustworthiness, for Set 1 we achieved $W = 0.7578$ ($p < 0.0001$), and for Set 2 we achieve $W = 0.7345$ ($p < 0.0001$). Therefore, there is a strong consensus between participants in terms of ranking different groups of advisors. The average ranking result shown in Figure 2 suggests that the relationship between review balances (from extremely negative to extremely positive) and perceived trustworthiness follows an inverted-U shape, and an extremely negative balanced review history is perceived as the most untrustworthy profile by buyers (2 versus 3.4 and 2.3 versus 2.95).





In the test regarding perceived expertise, for Set 1 we achieved $W = 0.2867$ ($p < 0.05$), and for Set 2 we achieve $W = 0.6451$ ($p < 0.0001$). This result indicates that the levels of consensus in Set 1 and Set 2 are weak and moderate, respectively. The averaged ranking result is shown in Figure 2, which suggests that perceived expertise does not increase linearly when review balance ranged from extremely positive to extremely negative. Meanwhile, participants' rankings about advisors with almost all negative reviews (Type I) are different (7.38 versus 3.38) across two sets.

In summary, the results from study 1 reject H1 because advisors with extremely negative review balance (Type I) were perceived to be low in expertise. H2 is supported, suggesting that advisors who always give the same ratings (either negative or positive) are not trustworthy to buyers.

Considering that the participants did not gain high consensus regarding the expertise of the advisors, it is interesting to further explore the influences of review balances on perceived credibility (especially expertise) of advisors.

Study 2

There are at least four issues in study 1, which limit the explanation power of the result. First, the sample is relatively small (20 participants). Second, the list of reviews only contains ratings, and it is not clear what the results would be when both ratings and comments are displayed (a real online review system usually displays both ratings and comments). Third, the measurements of trustworthiness and expertise are based on ranking, not on pre-validated questions. Ranking has its limitations, for example, it uses a one-to-one matching method between an advisor and a position and therefore, it might be difficult for participants to choose between two or more advisors when their trustworthiness/expertise perceived to be similar. Moreover, rankings only provide sequential data within a set but little is known about the differences across two sets. And fourth, the total number of reviews is not controlled.

The aim of study 2 is to further verify the results of study 1 by considering the limitations of study 1. First, a large sample was organized, including 200 participants; second,

both ratings and review comments were displayed to participants; third, pre-validated questions and Likert scale were used to measure participants' opinions. And fourth, perceived trustworthiness and expertise were evaluated in both high and low review count conditions.

Experiment preparation

To determine appropriate number of reviews in two conditions (high and low number of reviews), we manipulated five lists of advisors' review history, which contained 10, 40, 80, 120 and 200 reviews. We provided these review history lists to three Ph.D. students who were experienced online buyers. Their feedback suggested that 10 and 40 reviews could be treated as low number of reviews, but a list with only 10 reviews was usually not enough to form an attitude towards an advisor. Therefore, we set the value of low review number to 40. The feedback also suggested that a list with 200 reviews was beyond normal processing capacity, so we set the value of high review number to 200.

We built 10 advisors' review history lists based on selected 400 reviews. The details are shown in Table 3. We edited some of the reviews to make sure that these reviews did not conflict with each other. For example, one review may indicate that an advisor is a mother, but another review may indicate that the advisor is a father.

Details of experiment

We designed an online survey system which consisted of two parts: an advisor's review history and questions regarding trustworthiness and expertise. In the review history page, participants were told to imagine that they were shopping in Taobao as usual, and need to evaluate the credibility of an advisor. They should use the same amount of time to judge the advisor in our survey as in their regular purchase, and they could go to the questionnaire page as soon as they felt they have finished their judgment.

All questions in the survey were measured with 7-point Likert scale. Trustworthiness was measured by five items (dependable, honest, reliable, sincere and trustworthy); expertise was also measured by five items (expert, experienced, knowledgeable, qualified, skilled). These items were originally developed by Ohanian [25], and they have been adopted by many studies [49]. In order to do manipulation check, we used a question to ask participants to select one of the five conditions ($R < S$; $R < S$; $R \approx S$; $R > S$; $R > S$) which best fits what they see.

We invited 200 participants into our experiment. They were undergraduate students and they all had purchase experience in Taobao. Each participant was randomly

Table 3 Advisors' profile used in study 2

Type	Low review count (R,S)	High review count (R,S)
$R < S^a$	(1, 39)	(4, 196)
$R < S$	(12, 28)	(59, 141)
$R \approx S$	(19, 21)	(98, 102)
$R > S$	(29, 11)	(136, 64)
$R > S$	(40, 0)	(198, 2)

Note^a: R refers to number of positive ratings/reviews; S refers to number of negative ratings/reviews.

assigned into one of the ten conditions (5 types of review balance \times 2 types of review count). Therefore, each condition had 20 participants. This sample size provided an acceptable level of statistical power with an effective size of 0.50 at a two-tailed 5% significance level [50]. We selected undergraduate students as research subjects based on following two reasons: first, students provided an accessible sample when an experiment requires a large sample size [51]; second, young adults and university students are a typical group of online buyers, and similar sampling approach has also been employed in previous studies [52,51,17]. Moreover, a recent official survey shows that 56.4% of Chinese buyers in online marketplaces are aged between 20 and 29, 35.9% of consumers have (or are pursuing) bachelor degrees [41].

Analysis and result

All participants could correctly select the condition they were assigned to, indicating that our manipulations were successful. Table 4 shows the results of factor analysis (CFA) for both high and low review count conditions. All factor loadings were significant ($p < 0.01$), and ranged from 0.73 to 0.93. The composite reliability and Cronbach's alpha of each factor ranged from 0.86 to 0.94, demonstrating acceptable levels for internal reliability (the recommended threshold for these two indices is 0.7). All values of AVE shown in Table 4 are greater than the recommended value (0.5), suggesting that the latent constructs account for the majority of the variance in their indicators on average [53]. As a common rule, the presence of multi-collinearity issue is confirmed if Variance Inflation Factor (VIF) is higher than 10 [54]. More strictly, the VIF threshold of 3.3 has been recommended by Cenfetelli & Bassellier [55]. Table 4 shows that only two items (EXP2 and EXP3) from the high number reviews group are larger than 3.3 (but smaller than 10), indicating that multi-collinearity is not a serious issue.

We conducted two 5×2 ANOVA analyses on trustworthiness and expertise respectively. For trustworthiness, both review count ($F(1,190) = 4.045$, $p < 0.05$) and review balance conditions ($F(4,190) = 109.159$, $p < 0.001$) have significant main effects, but no significant interaction effect ($F(4,190) = 1.231$, $p > 0.05$). This result suggests that in general the participants perceived higher trustworthiness under the high review count

Table 4 Results from confirmation factor analysis in study 2

Constructs		Loading	C.R.	C.A.	AVE	VIF
Trustworthiness	TRU1	0.83/ 0.87 ^a	0.91/0.93	0.87/0.91	0.66/0.74	2.19/2.65
	TRU2	0.77/0.83				1.80/2.10
	TRU3	0.75/0.89				1.74/2.99
	TRU4	0.83/0.87				2.13/3.09
	TRU5	0.87/0.85				2.41/2.59
Expertise	EXP1	0.77/0.79	0.90/0.94	0.86/0.92	0.65/0.76	1.79/2.53
	EXP2	0.87/0.93				2.56/4.42
	EXP3	0.86/0.92				2.57/3.83
	EXP4	0.73/0.91				1.80/2.47
	EXP5	0.78/0.78				1.94/2.15

Note^a: the value on the left side of "/" is from the low number of reviews condition; the value on the right side of "/" is from the high number of reviews condition.

conditions than under low review count conditions (mean differences = 0.178, $p < 0.05$). And in both low and high review count conditions, the values of perceived trustworthiness are distributed in an inverted-U curve (see the repeated contrast of means shown in Table 5).

For expertise, both the review count ($F(1,190) = 5.656, p < 0.05$) and review balance conditions ($F(4, 190) = 35.906, p < 0.001$) have significant main effects. Moreover, a significant interaction effect is observed ($F(4, 190) = 13.05, p < 0.001$). This result suggests that the advisor’s expertise is perceived to be higher under low number of reviews condition than under high number of reviews condition (mean differences = 0.288, $p < 0.05$). And the values of perceived expertise are distributed differently across high and low number of reviews conditions. In low number of reviews condition, only the difference between means in conditions “R < <S” and “R < S” is negative (−0.36, but insignificant), suggesting that the perceived expertise linearly increases when review balance ranges from extremely positive to extremely negative. However, in the high number of reviews condition, the values of perceived expertise are distributed differently (an inverted-U shape). Especially when advisors have almost all negative reviews, they are perceived to be very low in expertise (see Table 5, repeated contrast of means between conditions “R < <S” and “R < S”: −2.64, $p < 0.001$).

In line with study 1, study 2 supports H2 but rejects H1. The results from both study 1 and study 2 show that buyers might misattribute low trustworthiness to low expertise, and this case might happen when buyers check an advisor who has a high number of reviews.

Study 3

The misattribution phenomenon found in study 1 and study 2 suggests that it is necessary to further explore the interplay between sub-dimensions of trust and expertise. Previous studies indicate that misattribution is usually a kind of affective response to a stimulus [56]. Similar to source credibility, trust is also a multifaceted variable, including both cognitive dimension and affective dimensions [22].

Cognitive trust is a kind of prediction based on people’s accumulated knowledge gained through observation of trustee’s behavior [22]. Affective trust is generated based on the positive emotions in the judgement process. Previous studies assume a positive impact of cognitive trust on affective trust because cognitive trust is a prerequisite for

Table 5 Means and repeated contrast results in study 2

Review balance	Perceived trustworthiness				Perceived expertise				
	Condition	Low count ^a	Repeated contrast ^b	High count	Repeated contrast	Low count	Repeated contrast	High count	Repeated contrast
R < <S		3.49 (0.72)	–	3.42 (1.06)	–	5.07 (0.56)	–	3.04 (0.73)	–
R < S		4.82 (0.66)	−1.33***	5.10 (0.54)	−1.68***	5.43 (0.57)	−0.36 ^{N.S.}	5.68 (0.73)	−2.64***
R ≈ S		5.37 (0.54)	−0.55*	5.79 (0.50)	−0.69*	4.58 (1.05)	0.85*	4.70 (1.07)	0.98**
R > S		6.09 (0.39)	−0.72**	6.39 (0.39)	−0.60*	4.40 (0.92)	0.18 ^{N.S.}	4.61 (1.01)	0.09 ^{N.S.}
R > >S		4.90 (0.57)	1.19***	4.86 (0.60)	1.53***	3.34 (0.81)	1.06**	3.35 (0.89)	1.26***

Note: ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; ^{N.S.}: $p > 0.05$; ^a: the values with parenthesis are standard deviations. ^b: The mean value in latter condition minus the mean value in former condition.

affective trust [57,22]. Cognitive trust has clear distinctions with expertise [7]. However, affective trust may have close relationship with expertise because of buyer's misattribution. Therefore, we conducted study 3 to explore the relationships among affective trust, cognitive trust and expertise in high number of reviews condition.

Details of experiment

A survey-based experiment was conducted. Detailed content of measurable items are shown in Table 6. Three measurable items (AFF3, AFF4, AFF5) for trust are extracted from previous study [57], while others are self-developed. Self-developed measures were used because no relevant items can be found in previous studies, and these items were developed to fit our research context well. Items used to measure expertise are extracted from Ohanian [25].

The experiment procedure is similar to the procedure in study 2. We invited 100 undergraduate students with Taobao purchase experience to take part in our experiment. The demographic information of participants is shown in Table 7. The number of participants meets the requirement of Partial Least Squares (PLS) analysis. Each participant was randomly assigned into one of the five review balance conditions with an advisors' review history containing 200 reviews. Participants were told to imagine that they were shopping in Taobao and need to judge the credibility of the advisor. Survey was provided as soon as the participants finished their judgement.

Table 6 Results of measurement model in study 3

Construct	Items	Content	C.R	C.A.	AVE	Loading	VIF
Cognitive trust	COG1	I see no reason to doubt his motivation to write reviews	0.94	0.91	0.75	0.77	2.26
	COG2	I think taking his review into consideration is a good decision				0.95	2.53
	COG3	I think I can rely on his reviews				0.77	2.79
	COG4	I think what he write in the reviews (pros and cons) is reasonable				0.92	2.79
	COG5	I think the review content and review activities make him a trustworthy advisor.				0.88	3.09
Affective trust	AFF1	I can feel his sincerity in writing reviews.	0.93	0.92	0.76	0.84	2.77
	AFF2	I am confident that he writes reviews based on his real experience.				0.95	3.51
	AFF3	I feel comfortable about relying on him for my purchase decision.				0.87	1.79
	AFF4	I feel secure about relying on him for my purchase decision				0.83	3.38
	AFF5	I feel content about relying on him for my purchase decision				0.85	3.64
Perceived expertise	EXP1	Expert-not an expert	0.93	0.91	0.74	0.88	2.94
	EXP2	Experienced-inexperienced				0.89	2.39
	EXP3	Knowledgeable-unknowledgeable				0.86	2.83
	EXP4	Qualified-unqualified				0.83	2.58
	EXP5	Skilled-unskilled				0.83	2.79

Note: S.D.: standard deviation. C.R.: Composite reliability. C.A.: Cronbach's alpha.

Table 7 Demographic information of participants

Items	Mean	S.D.	Min	Max	Comment
1. Age	22.24	1.11	19	25	
2. Gender	0.49	0.50	0 (female)	1 (male)	Male:49; Female:51
3. How much Taobao purchase experience do you have?	4.95	0.76	4	6	7 point scale (rarely-very frequently)

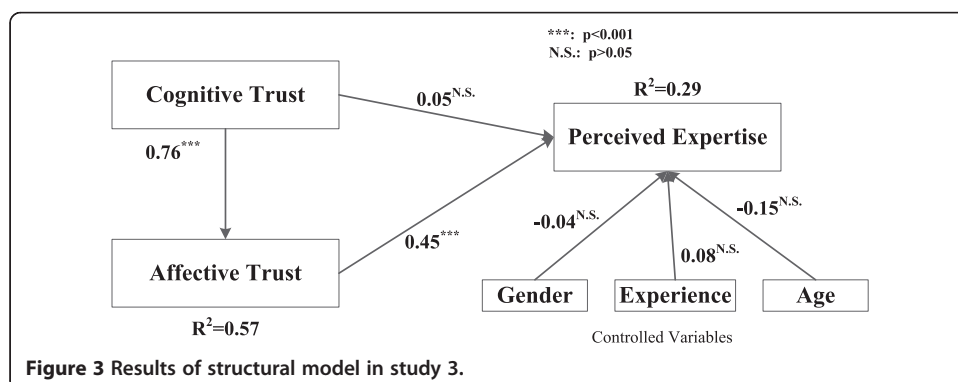
Structural equation modeling (SEM)-based PLS analysis was chosen to process survey data. This method was chosen in this study because it is suitable for exploratory study, and it requires neither large sample size nor multivariate normality of distribution [44]. We used WarpPLS 4.0 with bootstrapping to conduct PLS analysis. In line with other PLS softwares, the classic PLS algorithm was adopted.

Analysis and result

The analysis procedure is divided into two steps: test for measurement model and structural model. Table 6 shows the results of measurement model. All factor loadings were significant ($p < 0.001$), and ranged from 0.77 to 0.95. The composite reliability and Cronbach's alpha of each factor ranged from 0.91 to 0.94. All values of AVE are greater than 0.5. Finally, multi-collinearity is not a serious issue because the highest value of VIFs is only 3.64. These results indicate that our self-developed questions have good reliability and our survey data are suitable for further analysis.

In the test of structural model, first, age, gender and purchase experience are included as control variables. Results show that p values for these three variables are 0.06, 0.35 and 0.19. Therefore, no significant effects ($p > 0.05$) of control variables are found. As it is shown in Figure 3, the impact of cognitive trust on expertise is not significant ($\text{Beta} = 0.05, p > 0.05$). Cognitive trust has positive impact on affective trust ($\text{Beta} = 0.76, p < 0.001$), and affective trust positively influences expertise ($\text{Beta} = 0.45, p < 0.001$). The percentage of the variance explained (R^2) of affective trust and perceived expertise are 57% and 29%, indicating good explanation power.

The results of study 3 confirm the assumption of misattribution from trustworthiness to expertise, and further suggest that affective trust plays a significant role in determining expertise.



Summary and discussions

In online marketplaces, an advisor's credibility is important because buyers rely on advisor's reviews to make purchase decision. An advisor's profile is a major way for buyers to assess advisor's credibility. A profile usually includes identity-related information and review history. Disclosure of identity-related information has been found to be helpful in supporting buyers' judgment, however, the impacts of the review history remains unclear. In this research, we investigated the effects of review balance, an important aspect of review history. Study 1 investigated how buyers perceive advisors' trustworthiness and expertise based on different review balances. The results support H2 and show that perceived trustworthiness distributes in an inverted U-shaped curve when review balance ranges from extremely negative to extremely positive. Advisors with almost all positive or negative reviews are perceived to be not trustworthy, while advisors who write mixed reviews are perceived to be trustworthy. This result is in line with psychological studies [45], suggesting that mixed positive and negative reviews could enhance buyers favorable judgement towards a target (a seller, a product or an advisor). The finding is also supported by data mining studies, which treat advisors with all negative reviews as unusual cases with low trustworthiness [30,31,34,58,59]. An unexpected result in study 1 is that perceived expertise does not decrease linearly when review balance ranges from extremely negative to extremely positive. Therefore H1 is rejected.

An advisor with almost all positive reviews might be seen as an easy-to-satisfy buyer. As the proportion of negative reviews increases, advisor's perceived strictness on evaluating product increases. However, an advisor with extremely high proportion of negative reviews is perceived to be low in expertise. This result implies that buyers might misattribute low trustworthiness to low expertise. Many trust-related misattributions have been mentioned in previous studies. For example, alcoholism, drug abuse, and mental illness among managers can harm employee's trust towards the organization [60]; people with positive emotions (e.g., happiness and gratitude) are more inclined to trust than people with negative emotions (e.g., anger, sadness) [20]. This phenomenon occurs because affective states, even if they are caused by unrelated events, usually serve as an information aid in people's judgement.

Study 2 addresses some limitations of study 1 by incorporating larger sample size and more variables. The results of study 2 are consistent with those found in study 1, and again reject H1 but support H2. Study 2 further suggests that buyer's misattribution behavior is more likely to happen under high processing effort condition (advisor with high number of reviews). Currently there is little evidence to support the direct relationship between stress and misattribution. However, processing a high number of reviews can cause low processing fluency, which then leads to negative affective states [61].

The result of study 3 shows that expertise is positively related to affective trust, while not significantly related to cognitive trust. It provides evidence to explain why low trustworthiness leads to low expertise. Previous studies, however, neglect the affective aspect of trust and argue that trustworthiness and expertise are clearly distinguishable [7].

Implications

The results of this study yielded a couple of theoretical implications. First, previous studies on online marketplace mainly focus on the importance of advisors' review for

potential buyers in evaluating the trustworthiness of sellers. They advocate that the existence of inconsistent reviews, rather than majority positive or majority negative reviews, better reflects the seller's credibility. However, they did not consider different credibility of advisors. This study explores how the advisor's profile signals credibility meaning to buyers. Second, a large amount of work on advisors' credibility focuses on static personal information (e.g., gender, hobbies). This study moves a step further to evaluate the impact of review balance shown in review history. It is worthwhile to explore review history since it can provide valuable information to judge advisors' credibility. Third, this study enriches extant knowledge about the relationship between trustworthiness and expertise. Previous studies mention that people can easily distinguish between trustworthiness and expertise [7], and in some experiment, manipulations of trustworthiness and expertise were not found to influence each other [62]. However, in this study, low expertise is found to easily be misattributed from low trustworthiness, especially when buyers face advisors with a high number of reviews.

This study also generates practical implications for the design of mechanisms to support credibility judgement. First, there are many ways to assigning trust value when buyers and sellers are strangers, including initializing trust values based on beta distribution, and incorporating social network attributions. We argue that assigning trust values should take subjective perception into consideration. The results of this study could serve as a reference for assigning credibility values of advisors. Second, some trust models compute advisors' trustworthiness based on the degree of consensus among advisors [36]. Such method might not be suitable in a marketplace with a high proportion of fake positive reviews, because these models assume other advisors are credible. The results of this study could be helpful to refine existing trust models by reducing the importance of consensus in considering trustworthiness. For example, a malicious advisor with all positive reviews might be judged in existing models as highly trustworthy because his reviews are in agreement with others, however, he will be considered to be less trustworthy in revised trust model.

Limitations and future work

This study has five limitations, which affect the generalizability of our findings. First, although our research participants (mostly undergraduates) reflect a typical group of buyers in online marketplace, they cannot be representative of the whole consumer community. Moreover, our participants were required to have purchase experience and they were aware of unfair/review issues in online marketplace, therefore our findings cannot fully explain how new buyers perceive credibility of advisors with different review histories. We will extend our work by inviting participants with various backgrounds in future work. Second, in our experiment, an advisor's reviews for all sellers were listed together, but the differences (e.g., reputation) among different sellers were not considered. We argue that discarding sellers' difference does not significantly affect our result because it is unlikely for a buyer to further judge characteristics of sellers who are listed in an advisor's profile page. This issue will be considered in future work as a pretest before formal experiment. Third, our results cannot explain how buyers perceive an advisor who only has a few reviews. Buyers usually cannot make judgement based on a short review history list (e.g., only one or two reviews). Fourth, different

online marketplaces have different characteristics. Our target platform (Taobao) has serious unfair rating/review problem, while this issue might not be a problem in other platforms. Therefore, buyers in Taobao are assumed to have more knowledge about identifying advisors with low credibility. Fifth, in real purchase, buyers usually have to judge a list of advisors, while our experiments (study 2 and 3) only required participants to judge one advisor. The judgement of a list of advisor might be affected by the sequence of the list (primacy effect: buyers can only remember the credibility of the first advisor) and the information overload (e.g., buyers only judge a few advisors in the list). In future study, we will aim at measuring trust attitude towards a seller by providing buyers with a list of advisors.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

KW and ZN proposed initial idea of this paper. JV provided comments and helped KW and ZN with the formulation of research framework. KW and ZN conducted study 1 and KW conducted study 2 and 3. KW and ZN drafted the paper. IA proofread the draft and provided comments and suggestions. KW revised this paper according to reviewers' comments. JV supervised the revising process, proofread revised paper and confirmed responses to review comments. KW was in charge of submitting the paper and corresponding with the editors of the journal and Springer Open Production Team. All authors read and approved the final manuscript.

Acknowledgement

This work has been supported by NSERC through a Discovery Grant and a Discovery Accelerator Supplement Grant.

Received: 2 December 2014 Accepted: 12 January 2015

Published online: 12 February 2015

References

- Li M, Huang L, Tan C-H, Wei K-K (2013) Helpfulness of online product reviews as seen by consumers: Source and content features. *Int J Electron Commer* 17(4):101–136
- Lee J, Park D-H, Han I (2008) The effect of negative online consumer reviews on product attitude: An information processing view. *Electron Commer Res Appl* 7(3):341–352
- Park D-H, Lee J, Han I (2007) The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *Int J Electron Commer* 11(4):125–148
- Chua AYK, Banerjee S (2014) Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *Journal of the Association for Information Science and Technology*
- Zhang Y, Bian J, Zhu W (2012) Trust fraud: A crucial challenge for china's e-commerce market. *Electron Commer Res Appl* 12(5):299–308
- Dellarocas C (2000) Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In: *Proceedings of the 2nd ACM conference on Electronic commerce*. ACM, New York, pp 150–157
- Sprecker K (2002) How involvement, citation style, and funding source affect the credibility of university scientists. *Sci Commun* 24(1):72–97
- Chen Z, Yang J (2009) Credit fraud control and credit system optimization on c2c marketplaces. In: *Proceedings of the 42nd Hawaii International Conference on System Sciences*. Hawaii, HI, IEEE Computer Society Press, Washington
- Lim Y-S, Van Der Heide B (2014) Evaluating the wisdom of strangers: The perceived credibility of online consumer reviews on yelp. *Journal of Computer-Mediated Communication*
- Park H, Xiang Z, Josiam B, Kim H (2013) Personal profile information as cues of credibility in online travel reviews. *Anatolia* 25(1):13–23
- Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inf Syst Res* 19(3):291–313
- Ghose A, Ipeirotis PG (2011) Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Trans Knowl Data Eng* 23(10):1498–1512
- Heesacker M, Petty RE, Cacioppo JT (1983) Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *J Pers* 51(4):653–666
- Tintarev N, Masthoff J (2007) A survey of explanations in recommender systems. In: *IEEE 23rd International Conference on Data Engineering Workshop*. IEEE Computer Society Press, Washington, pp 801–810
- Rezaei S (2015) Segmenting consumer decision-making styles (cdms) toward marketing practice: A partial least squares (pls) path modeling approach. *J Retail Consum Serv* 22:1–15
- Mcallister DJ (1997) The second face of trust: Reflections on the dark side of interpersonal trust in organizations. *Res Negotiation Organ* 6:87–111
- Tamanidis T, Owusu-Frimpong N, Nwankwo S, Omar M (2015) Why we buy? Modeling consumer selection of referents. *J Retail Consum Serv* 22:24–36
- Park J, Gunn F, Han S-L (2012) Multidimensional trust building in e-retailing: Cross-cultural differences in trust formation and implications for perceived risk. *J Retail Consum Serv* 19(3):304–312

19. Willemsen LM, Neijens PC, Bronner F (2012) The ironic effect of source identification on the perceived credibility of online product reviewers. *J Comput-Mediat Commun* 18(1):16–31
20. Dunn JR, Schweitzer ME (2005) Feeling and believing: the influence of emotion on trust. *J Pers Soc Psychol* 88(5):736
21. Willemsen LM, Neijens PC, Bronner F, De Ridder JA (2011) "Highly recommended!" the content characteristics and perceived usefulness of online consumer reviews. *J Comput-Mediat Commun* 17(1):19–38
22. Johnson D, Grayson K (2005) Cognitive and affective trust in service relationships. *J Bus Res* 58(4):500–507
23. Heider F (2013) *The psychology of interpersonal relations*. Psychology Press, New Jersey
24. Pornpitakpan C (2004) The persuasiveness of source credibility: a critical review of five decades' evidence. *J Appl Soc Psychol* 34(2):243–281
25. Ohanian R (1990) Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *J Advert* 19(3):39–52
26. Sparks BA, Browning V (2011) The impact of online reviews on hotel booking intentions and perception of trust. *Tour Manag* 32(6):1310–1323
27. Huang L, Tan C-H, Ke W, Wei K-K (2014) Do we order product review information display? How? *Information & Management*
28. Khare A, Labrecque LI, Asare AK (2011) The assimilative and contrastive effects of word-of-mouth volume: An experimental examination of online consumer ratings. *J Retail* 87(1):111–126
29. Mudambi SM, Schuff D (2010) What makes a helpful online review? A study of customer reviews on amazon.com. *Manag Inf Syst Q* 34(1):185–200
30. Metzger MJ, Flanagin AJ, Medders RB (2010) Social and heuristic approaches to credibility evaluation online. *J Commun* 60(3):413–439
31. Chou SY, Picazo-Vela S, Pearson JM (2013) The effect of online review configurations, prices, and personality on online purchase decisions: a study of online review profiles on ebay. *J Internet Commer* 12(2):131–153
32. Amabile TM (1983) Brilliant but cruel: perceptions of negative evaluators. *J Exp Soc Psychol* 19(2):146–156
33. Ashforth BE, Mael F (1989) Social identity theory and the organization. *Acad Manag Rev* 14(1):20–39
34. Jindal N, Liu B, Lim E-P (2010) Finding unusual review patterns using unexpected rules. In: *Proceedings of the 19th ACM international conference on Information and knowledge management*. ACM, New York, pp 1549–1552
35. Ziegler C-N, Golbeck J (2007) Investigating interactions of trust and interest similarity. *Decis Support Syst* 43(2):460–475
36. Zhang J, Cohen R (2008) Evaluating the trustworthiness of advice about seller agents in e-marketplaces: a personalized approach. *Electron Commer Res Appl* 7(3):330–340
37. Zhu Y, Zhang W, Yu C (2011) Detection of feedback reputation fraud in taobao using social network theory. In: *2011 International Joint Conference on Service Sciences (IJCSS)*. IEEE Computer Society Press, Washington, pp 188–192
38. Popper B (2014) Alibaba has more sales than amazon and ebay combined, but will americans trust it? <http://www.theverge.com/2014/5/7/5690596/meet-alibaba-the-e-commerce-giant-with-more-sales-than-amazon-and-ebay>.
39. Clemes MD, Gan C, Zhang J (2014) An empirical analysis of online shopping adoption in beijing, china. *J Retail Consum Serv* 21(3):364–375
40. Moneyweek (2011) Gray industry cluster parasitism: Four complicated interest chains on taobao. <http://finance.stockstar.com/MS2011050300000899.shtml>.
41. Cnnic (2014) 2013 statistical report on chinese internet shopping https://www.cnnic.net.cn/hlwfzyj/hlwxzbg/dzswbg/201404/t20140421_46598.htm.
42. Sen S, Lerman D (2007) Why are you telling me this? An examination into negative consumer reviews on the web. *J Interact Mark* 21(4):76–94
43. Pavlou PA, Dimoka A (2006) The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Inf Syst Res* 17(4):392–414
44. Kallweit K, Spreer P, Toporowski W (2014) Why do customers use self-service information technologies in retail? The mediating effect of perceived service quality. *J Retail Consum Serv* 21(3):268–276
45. Ein-Gar D, Shiv B, Tormala ZL (2012) When blemishing leads to blossoming: the positive effect of negative information. *J Consum Res* 38(5):846–859
46. Berger J, Sorensen AT, Rasmussen SJ (2010) Positive effects of negative publicity: when negative reviews increase sales. *Mark Sci* 29(5):815–827
47. Xie H, Miao L, Kuo P-J, Lee B-Y (2011) Consumers' responses to ambivalent online hotel reviews: the role of perceived source credibility and pre-decisional disposition. *Int J Hosp Manag* 30(1):178–183
48. Nevo D, Chan YE (2007) A delphi study of knowledge management systems: scope and requirements. *Inform Manag* 44(6):583–597
49. Senecal S, Nantel J (2004) The influence of online product recommendations on consumers' online choices. *J Retail* 80(2):159–169
50. Cohen J (2013) *Statistical power analysis for the behavioral sciences*. Academic, Routledge
51. Liqiong D, Poole MS (2010) Affect in web interfaces: a study of the impacts of web page visual complexity and order. *MIS Q* 34(4):711–A710
52. Kim DJ, Ferrin DL, Rao HR (2008) A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decis Support Syst* 44(2):544–564
53. Mackenzie SB, Podsakoff PM, Podsakoff NP (2011) Construct measurement and validation procedures in mis and behavioral research: Integrating new and existing techniques. *MIS Q* 35(2):293–A295
54. Mason CH, Perreault WD Jr (1991) Collinearity, power, and interpretation of multiple regression analysis. *J Mark Res* 28(3):268–280
55. Cenfetelli RT, Bassellier G (2009) Interpretation of formative measurement in information systems research. *MIS Q* 33(4):7
56. Payne BK, Hall DL, Cameron CD, Bishara AJ (2010) A process model of affect misattribution. *Bulletin, Personality and Social Psychology*
57. Sun H (2010) Sellers' trust and continued use of online marketplaces. *J Assoc Inf Syst* 11(4):2

58. Mukherjee A, Liu B, Wang J, Glance N, Jindal N (2011) Detecting group review spam. In: Proceedings of the 20th international conference companion on World Wide Web. ACM, New York, pp 93–94
59. Lim E-P, Nguyen V-A, Jindal N, Liu B, Lauw HW (2010) Detecting product review spammers using rating behaviors. In: Proceedings of the 19th ACM international conference on Information and knowledge management. ACM, New York, pp 939–948
60. Speller JL (1989) Executives in crisis: Recognizing and managing the alcoholic, drug-addicted, or mentally ill executive. Jossey-Bass, San Francisco
61. Mosteller J, Donthu N, Eroglu S (2014) The fluent online shopping experience. *J Bus Res* 67(11):2486–2493
62. Wiener JL, Mowen JC (1986) Source credibility: on the independent effects of trust and expertise. *Adv Consum Res* 13(1):306–310

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
