View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Institutional Knowledge at Singapore Management University

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems | School of Information Systems

9-2010

Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures

Noi Sian KOH
Singapore Management University, noisian.koh.2006@smu.edu.sg

Nan HU
Singapore Management University, hunan@smu.edu.sg

Eric K. CLEMONS
University of Pennsylvania

DOI: https://doi.org/10.1016/j.elerap.2010.04.001

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

Part of the Computer Sciences Commons, and the E-Commerce Commons

Citation

KOH, Noi Sian; HU, Nan; and CLEMONS, Eric K.. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. (2010). *Electronic Commerce Research and Applications*. 9, (5), 374-385. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/212

This Journal Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Electronic Commerce Research and Applications xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures

Noi Sian Koh [a,*], Nan Hu [a], Eric K. Clemons [b]

[a] Singapore Management University, 80 Stamford Road, Singapore 178902, Singapore
[b] The Wharton School of the University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA 19104-6340, United States

ARTICLE INFO

Article history:
Received 8 February 2010
Accepted 3 April 2010
Available online xxxx

Keywords:
Online reviews
Movie ratings
Cross-cultural comparisons
Under-reporting bias

ABSTRACT

This paper investigates when the *reported average* of online ratings matches the *perceived average* assessment of the population as a whole, including the average assessments of both raters and non-raters. We apply behavioral theory to capture intentions in rating online movie reviews in two dissimilar countries – China and the United States. We argue that consumers' rating behaviors are affected by cultural influences and that they are influenced in predictable ways. Based on data collected from IMDB.com and Douban.com, we found significant differences across raters from these two different cultures. Additionally, we examined how cultural elements influence rating behavior for a hybrid culture – Singapore. To study whether online consumer reviews are subjected to under-reporting bias, which is, consumers with extreme opinions are more likely to report their opinions than consumers with moderate reviews causing online reviews to be a biased estimator of a product's true quality, we compare the consumer reviews posted online with those from an experimental study. Our results shows that under-reporting is more prevalent among US online network, thus online reviews are a better movie perceived quality proxy in China and Singapore than in the US.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

User-generated online reviews are a major source of information for movie-goers and can reduce product uncertainty and help consumers infer product quality. Virtually all models for monetizing online social networks, from the well known like Facebook and YouTube to the more obscure like FirstWivesWorld, are based on trust and shared social values (Clemons et al. 2007). The most successful, like the relationship between TripAdvisor and Hotels.com, are based on trust; the greatest failures, like Facebook's Beacon, occur when this trust is violated (Clemons et al. 2007). Still, consumers do pay attention to what has been written in online social networks to make their purchase decisions (Chatterjee 2001, Chevalier and Mayzlin 2006, Clemons 2008, Clemons and Gao 2008, Dellarocas 2003, Senecal and Nantel 2004). Prior research on consumer decision making has established that online reviews are considered more credible and trustworthy by consumers than traditional sources of information (Bickart and Schindler 2001, Li and Bernoff 2008). However, to what extent can the evaluations posted by individuals in online networks be considered reliable and representative of the general consensus? This is crucial to understanding the prospects for monetizing social networks, and even to their continued relevance in marketing.

Prior research (Hu et al. 2006) has found evidence that online reviews may not be representative of the general consensus opinions due to under-reporting bias. Under-reporting bias is a form of self-selection bias described in the literature on satisfaction (Anderson 1998). Consumers who are very satisfied or very dissatisfied will be more motivated to voice their opinions through reviews and thus are more likely to do so. It has been found that under-reporting bias does exist in certain US online review websites (Hu et al. 2006). The *average of reported quality ratings* (created by a small population of those sufficiently motivated to post their reviews) do not match the *average of perceived quality assessments* of the general population. Since consumers are becoming increasingly dependent on online reviews to make purchase decisions, we studied raters' behaviors to reveal whether under-reporting bias exists across cultures, and whether online consumer rating behavior will yield biased or unbiased estimators of a product's quality in various markets.[1]

Since each posted online review is an assessment of a single individual's perceived quality of a product, this study first explores how such reported quality could be influenced by cultural factors.

* Corresponding author. Tel.: +65 91918362; fax:+65 68280468.
E-mail addresses: noisian.koh.2006@smu.edu.sg (N.S. Koh), hunan@smu.edu.sg (N. Hu), clemons@wharton.upenn.edu (E.K. Clemons).

[1] In this paper, we will be using the terms 'rater', 'reviewer' and 'consumer' interchangeably.

1567-4223/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.elerap.2010.04.001

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

The behavior of individuals in online networks can be very different, may vary in systematic ways across cultures, and may differ from offline behavior as well.

Behavioral theory in social psychology asserts that specific salient beliefs influence behavioral intentions and subsequent behavior (Ajzen 1985, 1988, 1991). Employing constructs from behavioral theory—attitude, social norms and motivation—we seek to understand the following important questions, which to the best of our knowledge have not been answered in previous online review literatures:

- What factors motivate consumers to write online reviews?
- How does culture influence raters' behavior when writing reviews and how are cultural differences manifested in differences among ratings?

To identify the potential under-reporting bias that might render the mean of online movie reviews a biased estimator of movie quality, we compare the distribution of voluntarily posted online movie reviews to those reviews that we believe are closer to the distributions of true perceived quality. In the case of Chinese reviewers we were able to conduct a survey, in which respondents were asked to report their ratings for a number of movies that they have viewed and under what circumstances they are more likely to write online reviews. In the case of American reviewers, we could not get an adequate number of responses for the survey experiment. Thus, we replaced the survey instrument with a database of reviews collected from MovieLens. As raters of MovieLens are advised to rate as many movie titles as they can in order to get a good movie recommendation, we believe that, in MovieLens, the under-reporting bias is less severe. Comparing survey results to posted online reviews in a Chinese cultural environment, and comparing unbiased recommender reviews to posted online reviews in an American environment together show that under-reporting bias varies across different cultures—online reviews more accurately reflect a movie's perceived quality in Chinese online networks than in the US.

The rest of the paper is organized as follows. The next section reviews related work. Section 3 describes the theoretical framework and our research hypotheses. The research setting and methodology are presented in Section 4. Sections 5 and 6 provide the analysis of our empirical findings. Section 7 provides a robustness check on our findings. Section 8 discusses the limitations and concludes with suggestions for future research.

2. Related work

To understand the behavior of movie raters from different cultures, we draw upon some of the behavioral theory on attitude, social norms, and under-reporting bias and examine how their behaviors are influenced by cultural differences. This section covers the work relating to attitude, social norms, cultural differences and under-reporting biases.

2.1. Attitude and social norms

Taylor and Todd (1995) describe a construct that argues that behavioral beliefs influence attitudes, which in turn determine intentions and actual behavior. Behavioral beliefs arising from social pressure are termed normative beliefs (Ajzen 1991), also termed *social norms*, the influence created by a person's normative beliefs that others approve or disapprove a particular behavior. People's intentions to perform a particular action are influenced by social norms, or by their perception that important others think they ought to perform those actions. In our context, social norms refer to the influence from consumers' normative belief that the behavior is accepted, encouraged, and promoted by their social circle. Consumers may believe that their family, friends, and even online peers would favor certain online opinions, and this belief tends to influence their intentions and opinions. We examine how offline interactions and social norms influence online social network behavior.

2.2. Cultural differences

Hofstede's (1980) cultural dimensions serve as the most influential theory of culture and cultural differences in research in the social sciences (Nakata and Sivakumar 2001); his categorization of national societies is also widely used as the basis of applied research in the study of marketing differences across cultures and in e-commerce studies (Pavlou and Lin 2002). His cultural framework has also received strong empirical support (Sondergaard 1994). The framework was generated through the most extensive examination of cross-national values ever undertaken, involving 116,000 respondents from 40 countries (Pavlou and Lin 2002). The results were consistent with the findings in 38 other studies (Nakata and Sivakumar 2001). Hofstede separated cultures on the basis of (a) masculinity–femininity, (b) individualism–collectivism, (c) power distance, (d) uncertainty avoidance, and the recent addition of the Confucian dimension of (e) long-term orientation (Hofstede 2001). Our work starts by accepting Hofstede's framework; we are not testing it for validity, but attempting to demonstrate whether the behavior of online raters is consistent with this theory. We focus on the implications of dimensions (b) individualism–collectivism and (e) long-term orientation.

Individualism–collectivism refers to the basic level of behavior regulation of either individuals or groups. Individualists view self and immediate family as relatively more important than the collective. *Long-term orientation* as described by Hofstede suggests following tradition, perseverance and the practice of benevolence; *short-term orientation* on the other hand, is the tendency towards consumption and materialism. As these are long-established and influential theories of culture and cultural differences, we will be using these cultural constructs in our conceptual development to better understand how different forms of national culture manifest themselves in online interactions (Ess and Sudweeks 2005). We caution the reader not to view these dimensions as merely cultural stereotypes. Hofstede is not suggesting that all Chinese are benevolent towards all other humans, or even towards all other Chinese, in their online behavior, nor is he suggesting that Western culture is without benevolence and the Golden Rule of "Do unto others..." Hofstede is suggesting that with a large enough sample, differences in cultural norms are readily observable.

2.3. Under-reporting biases

Hu et al. (2006) found evidence of two self-selection biases, acquisition bias and under-reporting bias, in the reporting of online consumer reviews, both of which render mean ratings a biased estimator of product quality. Acquisition bias refers to the situation that only consumers with a favorable disposition towards a product will acquire the product. Since only consumers with a pre-acquisition utility perception higher than the product's posted price are willing to pay the price to acquire, and thus have the chance to review the product, this creates a bias towards a greater number of positive product reviews. Second, consumers who are greatly satisfied or greatly dissatisfied are more likely to report their reviews; correspondingly, those consumers with more moderate sentiments are less likely to post a review. This is termed *under-reporting bias*, since it applies to consumers who have acquired

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

the product and specifically to which consumers are more likely or less likely to report.

Based on the data collected from Amazon.com and an offline survey conducted on US customers, Hu et al. (2006) documented that while online consumer reviews have a U-shaped distribution, actual consumer assessments for the same set of products are normally distributed. They concluded that online reviews do not reflect a product's perceived quality across the population of all users, which they term its *true perceived quality*. Rather, online reviews quite naturally reflect the views of those who post them, which differ from true perceived quality because of under-reporting by those customers with moderate views.

However, their study did not examine the raters' attitude and social norms across different cultures. Since the degree of under-reporting bias might vary across cultures, we set out to understand raters' behavior across different cultures—American and Chinese—and to identify under what circumstances online reviews might or might not reflect a product's true perceived quality in different settings. Since a large proportion of the movies watched by Chinese consumers were downloaded at no cost, acquisition bias is not very significant in our context. Hence, we focused on under-reporting bias for this study.

The United States and China were chosen for this study because they represent almost reverse positions on several important cultural dimensions (Hofstede 1980). In addition, we chose to collect data on Singapore because of its mixture of Western and Eastern culture, which allows us to see how culture mediates attitudes and behaviors in such a hybrid culture.

3. Conceptual development

The proposed research model of online movie reviewers' behavior is adapted from Pavlou and Lin (2002). The dependent variable, online review behavior, is measured by the rating assessment each reviewer gives to each movie, and captures consumers' reviewing behaviors. Drawing from behavioral theory in social psychology, two factors that directly influence reviewers' intentions towards the reviewing process included in this study are attitudes towards the movie and social norms regarding what is customary in rating behavior (Ajzen 1991, Pavlou 2002). We investigate the relationships among these in terms of cultural differences, using on the dimensions of individualism/collectivism and long-term/short-term orientation. In addition, we gathered survey data that enabled us to look at the motivation for consumers to write online reviews.

3.1. Attitude

Attitude has been used as a predictive factor that influences behavioral intention in multiple theories, such as the *theory of planned behavior* (TPB) (Ajzen 1988) and the *theory of reasoned action* (TRA) (Fishbein and Ajzen 1975). These theories have gained substantial empirical support (Madden et al. 1992, Pavlou and Lin 2002). Attitude here refers to an overall evaluation of the movie that an individual has viewed. A favorable attitude towards a movie will positively influence the rating of online movie reviews. Of course, it is not surprising that a reviewer's attitude towards a movie, which is essentially his or her assessment of that movie, will be a prime determinant of the content of the review posted. But it is not the only determinant.

3.2. Social norms

Social influence is related to Hofstede's dimension of individualism/collectivism, and is a second factor that directly influences online reviews. Collectivism refers to the extent to which individuals feel themselves to be integrated into groups and the extent to which opinions are informed by group norms and expectations or even formed based on these norms and expectations (Hofstede and Bond 1988). Members of individualistic societies are more likely to value freedom of expression, while those of collectivistic culture are more likely to seek group consensus. China has for centuries been highly collectivist; in particular, Chinese attention to group norms predates collectivism in the sense imposed by Communism and indeed goes back to China's Confucian heritage. Conversely, the United States is among the most highly individualistic societies. Consequently, we expect there to be differences in the effect of societal influence on individual behavior, and specifically we expect to observe these differences by comparing online movie reviews contributed by members of the two cultures.

Collectivist societies have strong relations within the extended family and among friends and acquaintances (Hofstede and Bond 1988). Their group relations seek to maintain harmony by going along with the group's wishes and by promoting and maintaining long-term relationships (Bond and Smith 1996). We anticipate that members of a collectivist culture, such as China, would want to maintain harmonious relationships among participants, both as readers and as writers, in the online movie review website. On the other hand, we expect that US movie reviewers value freedom of expression more strongly and hence feel themselves to be more free to openly express their appreciation or great dissatisfaction of the movies they have viewed. Indeed, as noted by Jaron Lanier (2010) in his recent book, the anonymity made possible by websites seems to encourage the emergence of Internet trolls in the west and a practice he calls drive-by anonymous insults.

Due to their recent colonial history, Singaporeans have been influenced by Western culture, but because of their earlier history of Eastern culture, collectivist values are also strong. Values such as obedience and harmony are important, and they value intense friendships and trust within the family. Therefore, we anticipate that Singapore is more of a collectivist society than the United States. Since China is highly collectivist and the US is highly individualistic, we expect that the analysis of review patterns from China, Singapore, and the United States will reveal significant differences in reviewer behavior, consistent with Hofstede's cultural classification. We anticipate that the ratings posted by American reviewers will more clearly express their likes and dislikes for movies. On the other hand, the ratings given by Chinese reviewers will be more constrained and more narrowly confined within a tight range centered around the average of the ratings given by previous customers. Thus, we expect that an attitudinal difference in the reviewing intentions among the three countries, again consistent with Hofstede, as the following hypotheses propose.

Hypothesis 1 (**Collectivist Societies**). *Collectivist societies tend to place greater focus on harmony and thus tend to write fewer extremely negative reviews than individualist societies.*

Hypothesis 2 (**Freedom of Expression**). *The value placed upon freedom of expression is reflected more in the ratings of movies from individualist societies than those from collectivist societies.*

Hypothesis 3 (**Societal Norms**). *Societal norms have greater effect on the ratings of movies in collectivist societies than in individualist societies.*

According to Hofstede (2001), China is ranked extremely high on the dimension of long-term orientation, which reflects the impact of the teachings of Confucius on Chinese culture and society. One of the key principles of Confucian teaching is the basic human benevolence toward others and this consists of

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

treating others as one would like to be treated.[2] We therefore expect Chinese to be less willing to fully express their dislike in their ratings for bottom-ranked movies than American reviewers would be. Notice, we cannot distinguish whether Chinese have more generous views of the movies, or merely restrict themselves to more generous public statements and posted reviews, solely on the basis of the posted reviews. Additional hypotheses address these differences, and our methods for studying these differences are described in the section on research methodology. But, regardless of the motivation, just as we would expect Chinese reviewers to be more generous in general (as expressed in the Collectivist Societies Hypothesis, H1), we would expect them to be more generous even in the case of the worst and most disappointing experiences. Thus, we assert:

Hypothesis 4 (**Bottom-Ranked Movies**). *The ratings for bottom-ranked movies given by collectivist societies will be less extreme than in individualist societies.*

We have applied Hofstede's (2001) cultural classifications, allowing us to predict certain differences in behaviors across cultures, and our hypotheses allow us to analyze the extent to which the predicted differences do or do not appear as in the specific context of online behavior in the specific domain of online movie reviewing. We are not attempting to test Hofstede's theory, and we do not argue that all Chinese are collectivist, or that all Westerners are extreme individualists. We simply analyze millions of movie reviews posted on websites that cater primarily to Chinese or to American reviewers, and look for the predicted differences in rating behavior. The differences are indeed consistent with predictions based on Hofstede's theory. We understand that not all movie-goers rate the movies they have seen; significantly, the absence of ratings follows a pattern, with more extreme under-reporting bias in the United States than in China, which is consistent with our hypotheses and with Hofstede's cultural classification upon which they are based.

3.3. Motivation

Writing reviews seems to address basic human needs both for belonging to and gaining acceptance from groups in which they participate and for achieving status and recognition (Maslow 1943). We hypothesize that consumers who participate in the writing of online reviews are motivated to meet these needs. If this were true, then writing reviews would be based both on individual motivations and on the interaction of these motivations with social norms. An individual from a highly collectivist society would most definitely not achieve his desire to feel as if he were part of a group if his reviews violated the norms of the group, and would not receive self-esteem and recognition if his reviews were rejected because they violated the norms of the group. Thus, we assert the following two hypotheses:

Hypothesis 5 (**Social Needs**). *The motivation to write movie review is affected by people's social needs to feel a sense of belonging and sharing, which may require that reviews adhere to social norms.*

Hypotheses 6 (**Esteem and Recognition**). *The motivation to write movie review is affected by people's needs for esteem and recognition, which may require that reviews adhere to social norms.*

4. Research methodology

Our research methodology involves three specific sets of analyses. First, we compare the rating behavior of Chinese and American reviewer using data collected from Douban.com and IMDB.com. Second, we perform attitudinal studies to determine, to what extent the online reviews reflect a product's true perceived quality. This is described in Section 4.2. In addition, in Section 5, we study how likely a US or Chinese reviewer will be affected by the reviews that were posted previously. Finally, in Section 6, we investigate to what extent Singapore movie raters resemble those of China and to what extent they resemble those of the United States. We chose to perform these analyses only after observing differences between the US and Chinese movie raters' behavior.

4.1. Online review data

To study the cultural differences in online movie review behavior, we gathered the reviews from two online movie review websites, IMDB.com and Douban.com. IMDB.com was chosen because it is the largest online movie review website with over 57 million visitors each month. For the Chinese website, we chose Douban.com because it is a cloned version of IMDB in China, and is consistently ranked as one of the most popular online review website in China by Alexa Internet (2008). To ensure we are comparing members of the US and Chinese cultural communities, we only crawled the ratings from IMDB that were posted by US reviewers.[3]

Our data sample contains two datasets. For the first dataset, reviews were collected on 1000 movies randomly selected from IMDB and 1000 movies randomly selected from Douban, using a random counter on the movie identification number. While we expect that 1000 movies is a well representation of the movies across the two websites, it is possible that our results might be influenced by the differences between the movies selected from these websites.[4] To control for this, we conducted experiments with a second set of data. We first chose the Top 100 and Bottom 100-ranked movies in IMDB. Then based on these movie titles, we collected the same movies titles in Douban.com and their corresponding movie reviews from the two sites. By using the same movies to compare the rating pattern, we attempt to ensure that the observed differences in reviews are due to inherent differences in reviewer behavior rather than differences in the movies selected for comparison. The reason for focusing on those top and bottom-ranked movies is that if indeed under-reporting bias were present, we believed that it would be more likely to be observed for movies within such categories.

Each dataset has its own advantages and disadvantages. Using 1000 movies selected separately and at random from these two movies reviewing sites revealed in general how movie raters' behavior is different across these two websites. However, this introduces the possibility that the observed difference is driven

[2] While there is no immediately obvious connection to a Western observer between benevolence and a long-term orientation, historically Confucian teachings have stressed long-term orientation, collectivist ties to family and society, and a higher degree of benevolence within groups. Likewise, Hofstede describes group averages; at no point does he suggest that absence of benevolence in individualistic societies, or the absence of altruistic behavior in the West. Explicitly, the Judeo–Christian traditions of the West do acknowledge the importance of treating others as you would wish to be treated; still, Hofstede expects to see a greater degree of benevolence in the west, and, assuming this is true, we would expect to see differences between Chinese and American online movie reviewing.

[3] Each movie has a webpage that shows the ratings given by US raters. For instance, the ratings given by US viewers for the movie "The Godfather" can be obtained from http://www.imdb.com/title/tt0068646/ratings-usa. Nevertheless, the fact that a reviewer resides in the US does not necessarily mean that the reviewer is an American, or that the reviewer has adopted the behaviors that Hofstede typically associates with an individualist culture. However, to the extent that our data might include Chinese residing in the United States, this inclusion actually should lessen the strength of the effects we were measuring. Thus, our results are a conservative test and may actually understate the effect that we have claimed.

[4] The 1000 randomly selected movies in IMDB and 1000 randomly selected movies in Douban may be different.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

by the different move titles selected from the two sites. Using the top and bottom-ranked movies and then comparing reviews from the two sites eliminates the above possibility, but introduces the possibility that audience response to different movies might be different between the two cultures. For robustness check, we also collected data from the Top 100 and Bottom 100-ranked movies in Douban, and the same movie titles in IMDB and the corresponding movie reviews. This will be presented in Section 7.

The summary statistics of the dataset from both websites are shown in Tables 1 and 2.[5] Data collection for both batches started on December 20, 2008 and ended on January 15, 2009. For each item, we collected the movie title, movie ID and review information. Specifically, for each movie review, we gathered the numeric rating, review date and the original text of the review. On Douban, consumers can report an integer movie review on a 5-point Likert-type scale, anchored at 1-star = least satisfied and 5-stars = most satisfied. On IMDB, consumers can report an integer movie review on a 10-point Likert-type scale, anchored at 1-star = least satisfied and 10-stars = most satisfied.

Table 1
Statistics of dataset from IMDB.com.

Category	Number of movies	Number of ratings
Top 250	(Top) 100	2944,037
Bottom 100	100	191,411
Entire collection	(Random) 1000	691,739

Table 2
Statistics of dataset from Douban.com.

Category	Number of movies	Number of ratings
Top 250	(Top) 100	483,680
Bottom 100	100	4151
Entire collection	(Random) 1000	14,645,654

4.2. Experimental calibration with survey data

After comparing the rating distribution of Chinese consumers to that of US consumers, we conducted a survey in which respondents were asked to review the movies that they have viewed. Then we compared the survey results with those observed on IMDB and Douban. We expected that this survey mechanism would result in more balanced reviews with less under-reporting bias, and thus in ratings that more accurately reflect the community's perceived average quality for movies in our sample. This is essential to exploring our hypotheses about cultural differences in reviewing behavior.

Our respondents were university students attending business, information systems and economic courses. Each student was asked to review 16 movies that vary in terms of category and genre. For each movie, the subjects were asked to rate the movies they have viewed and report their intention and motivation to write online movie reviews. The online survey instrument was emailed to 1500 students composed of native Chinese speakers and Singaporeans who spoke Chinese fluently. Invitation emails explained the purpose of the study and requested participation. The respondents who clicked on the URL link provided in the email message were then directed to a website to take the online survey, of which 87 Chinese students and 212 Singaporeans responded. Participation was voluntary and the response rate was approximately 20%. The movie rating scales for the survey were based on those of Douban.com. A preliminary version of the survey was generated and reviewed by doctoral students for clarity. Finally, to verify the appropriateness of the survey, it was pre-tested with multiple research students who varied with age, gender and education. Since this attitudinal survey involved both Chinese and Singaporean nationals, we are able to assess the extent to which Singaporeans do or do not differ from Chinese in terms of their rating behaviors.

For US users, we employed ratings from MovieLens in place of our survey, because we were not able to get sufficient response for our survey among American movie reviewers.[6] Unlike IMDB, MovieLens is not an online review website but a recommender system. In MovieLens, users rate as many movies as possible (each user has to rate at least 35 movies) to improve the accuracy of the recommendations each will receive from the recommender system. Since in MovieLens the aim of the users for rating movies is to improve the quality of their own received recommendations, they are more likely to rate all the movies that they have watched rather than merely evaluating those to which they have had the most extreme responses. Hence, we believe that there is less under-reporting bias for reviews in MovieLens, which we believe will more accurately reflects the community's perceived quality for the movies in our sample.

5. Data analysis and results

5.1. Graphical data analysis and simple comparisons

We retrieved movie reviews of 1000 randomly selected movie titles from both Douban and IMDB, and we focused on movies with an average review of 3-stars in Douban and those with 5-stars in IMDB, the median ratings for the two sites. After that, we plotted the distribution of the ratings from Douban and from IMDB (as shown in Figs. 1 and 2). Theoretically movies with these median average ratings are more likely to be normally distributed, therefore, we expect to observe a normal distribution for both but as is evident from even a quick visual inspection of Figs. 1 and 2, we did not. The behaviors of the two populations indeed do differ.

Figs. 1 and 2 show that for all movies with average review score equal to the median (92 movies on IMDB and 151 on Douban), the rating histogram for IMDB is W-shaped, whereas for Douban the histogram is indeed unimodal and bell-shaped. Thus there are differences in the rating pattern even for average-ranked movies. To ensure that our results are not driven by the efforts of reviewers,

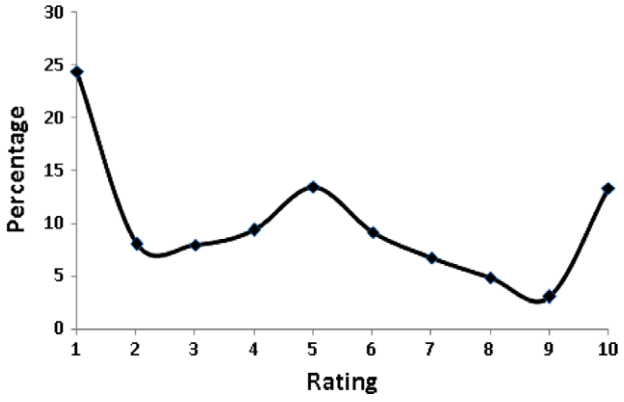

Fig. 1. IMDB movies with average rating = 5 (92 out of 1000 movie items).

[5] The data analysis is based on data collected prior to January, 15, 2009. The collection of our data was slow because we were constantly blocked by IMDB and Douban.

[6] These data were gathered for a GroupLens Research Project by the Department of Computer Science and Engineering at the University of Minnesota. See http://www.grouplens.org/taxonomy/term/14.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

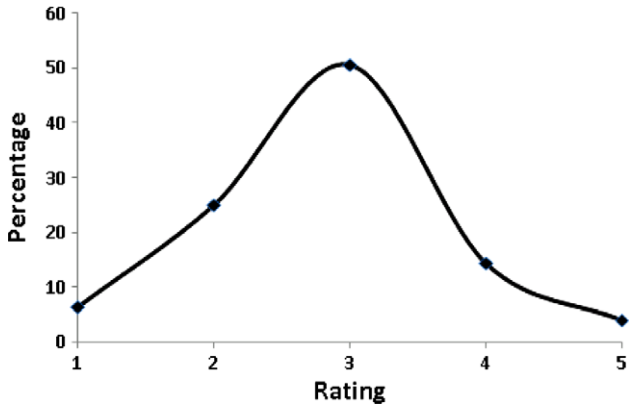

Fig. 2. Douban movies with average rating = 3 (151 out of 1000 movie items).

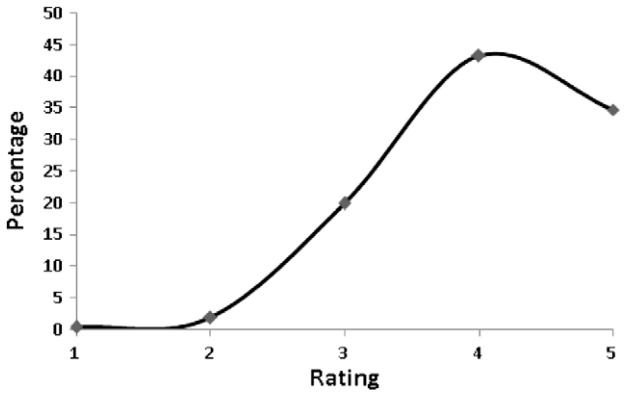

Fig. 4. Corresponding (top-ranked) movie ratings in Douban.

we checked if reviews with only ratings differ from reviews with both ratings and text comments. Since preparing a review is more time-consuming than merely providing a numeric evaluation and might suggest a more serious effort to assess accurately, we have compared the numeric ratings of reviews with and without textual reviews across IMDB and Douban and found them to be similar (Wu and Huberman 2007). Our results show that the rating histogram for Douban is bell-shaped, while that of IMDB is W-shaped.

In the work that follows, we compare individual rating behavior using the second dataset, in which we ensure that ratings are for the same set of movies. Since the movies being compared are the same, any observed differences in the rating patterns are driven by the cultural differences between the rating populations, and not by the selection of movies included in each sample. We focused on ratings of the most extreme movies, the top-ranked and bottom-ranked movies, in order to examine the most extreme rating behavior. Unless we state otherwise, all subsequent analysis was done using the second dataset.

Figs. 3 and 4 show the rating distributions for the top-ranked movies in IMDB and the corresponding rating distribution for the same set of movies in Douban. Fig. 3 once again displays the characteristic W-shaped distribution that we previously observed among US raters, whereas Fig. 4 once again has a unimodal distribution. It seems that on average, Chinese reviewers are more reserved in giving the highest ratings than the US reviewers; they are also more reluctant to assign the most negative reviews. From both figures, we see that on the whole, the rating behavior of reviewers of top-ranked movies were not very different between both cultures, except for some modest limitation of top and bottom reviews among Chinese reviewers.

Figs. 5 and 6 show the distribution of the ratings for the Bottom 100 movies in IMDB and the corresponding rating distribution for

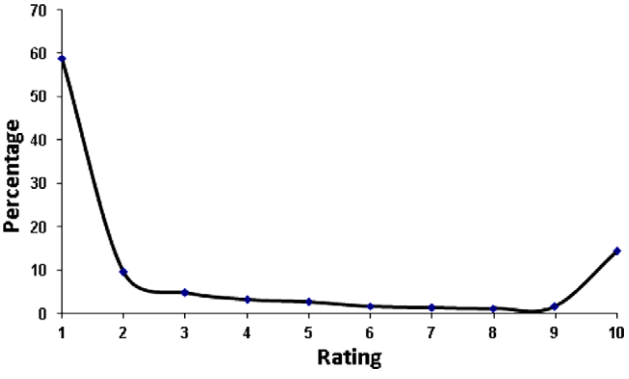

Fig. 5. IMDB bottom-ranked Movie Ratings.

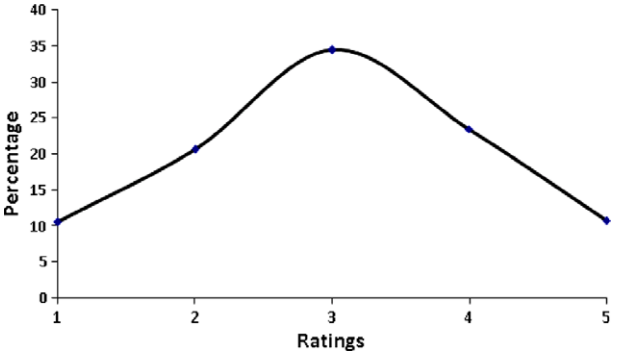

Fig. 6. Corresponding (bottom-ranked) movie ratings in Douban.

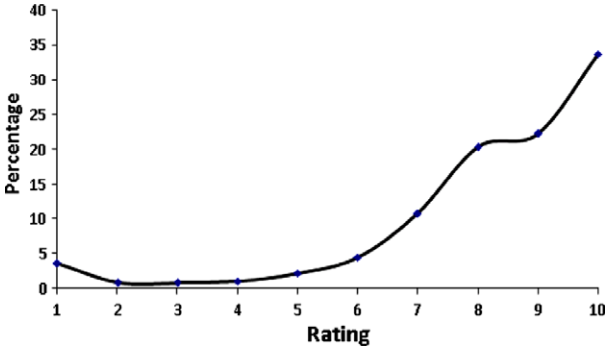

Fig. 3. IMDB top-ranked movie ratings.

the same set of movies in Douban. Unlike the results for the top-ranked movies, there is a great difference between Chinese and American reviews for IMDB bottom-ranked movies. Fig. 5 shows that for the US reviewers, the largest number of reviewers gave a very low rating of 1-star, resulting in a U-shaped distribution. For the Chinese reviewers, even when a movie is bad, the online reviews still demonstrate a bell-shaped distribution, as shown in Fig. 6.[7] This supports the Bottom-Ranked Movies Hypothesis (H4), which argues that the ratings for bad movies given in collectivist societies will be less extreme than those in individualist societies. US reviewers on the other hand are more inclined to express their

[7] To check the consistency of this result, we conducted the same analysis based on independent datasets collected prior to September 15 and prior to December 15, 2008. We still find the W-shaped distribution for the US and a bell-shaped distribution for the Chinese.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

dissatisfaction vigorously and openly, and often in the most extreme terms, which is consistent with the Freedom of Expression Hypothesis (H2). This probably partly explains the difference in the number of reviews for the Bottom 100 movie category in IMDB (191,411 total ratings) and Douban (4151 total ratings), providing support for the Collectivist Societies Hypothesis (H1). (A competing explanation is that the very worst movies screened in America are unlikely to be watched by Chinese audiences, which may also explain some of the observed difference in counts of reviews for the worst ranked movies.)

Prior studies (e.g., Tourangeau et al. 2000, Poulton 1989), have reported that respondents tend to avoid the extremes in surveys. In psychophysics, this trend is called response contraction bias (Poulton 1989). However, this is not what we have observed in IMDB.com; there are in fact more responses at the extreme endpoints of 1s and 10s, this is the case even when we convert the results to the same scale as shown in Figs. 7 and 8. For Figs. 3–6, if we first convert the 10-point IMDB scale to a 5-point scale, consistent with Douban, and then remove the extreme ratings—the ratings of 1 and 5—this will result in Figs. 9–12. The results suggest that the extremes are over-represented in the American rater population relative to the Chinese rater population, and perhaps the extremes in the American rater population are even over-represented relative to the American population of movie-goers more generally. There are several explanations. Perhaps Americans are more honest and willing to post extreme views because they are less influenced by the mean. Another explanation is that Americans might try to be different by giving extreme ratings, since simply giving an average rating does not show that they are individuals. Alternatively, Americans may be less willing even to rate unless extremely motivated by very strong attitudes, positive or negative, towards the

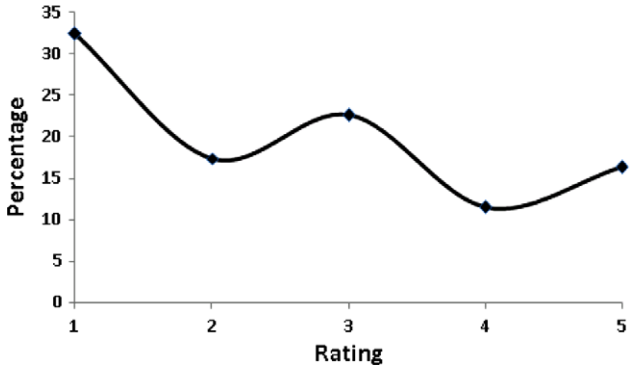

Fig. 7. IMDB movies with average rating (92 out of 1000 movie items) on same scale.

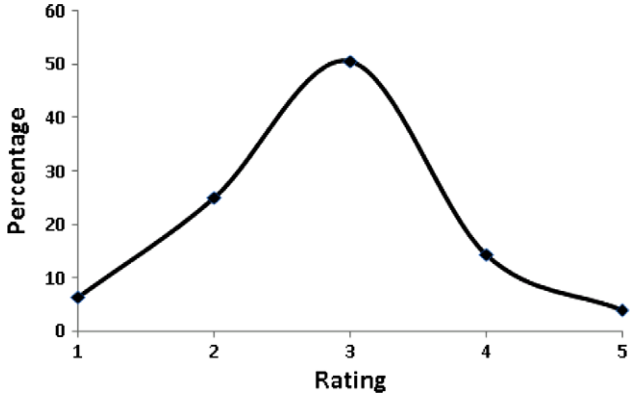

Fig. 8. Douban movies with average rating (151 out of 1000 movie items) on same scale.

Fig. 9. IMDB top-ranked without extremes.

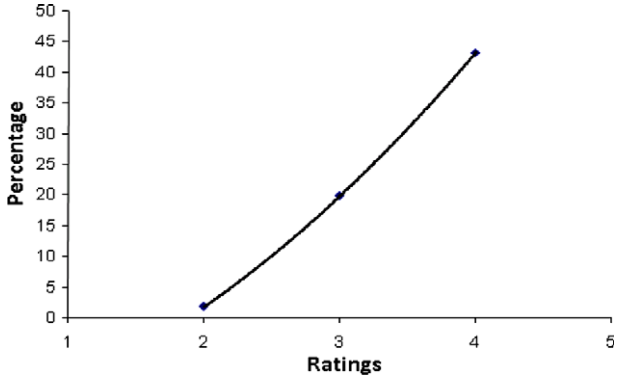

Fig. 10. Corresponding top-ranked without extremes in Douban.

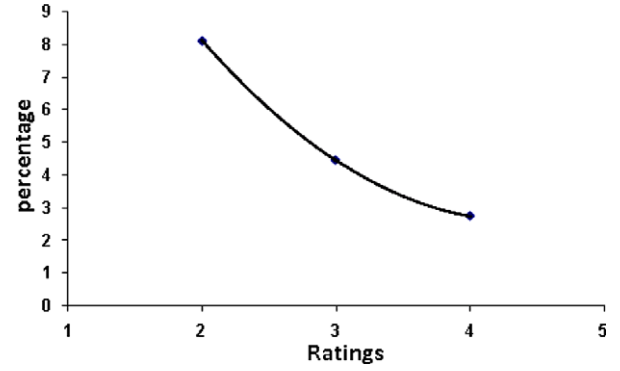

Fig. 11. IMDB bottom-ranked without extremes.

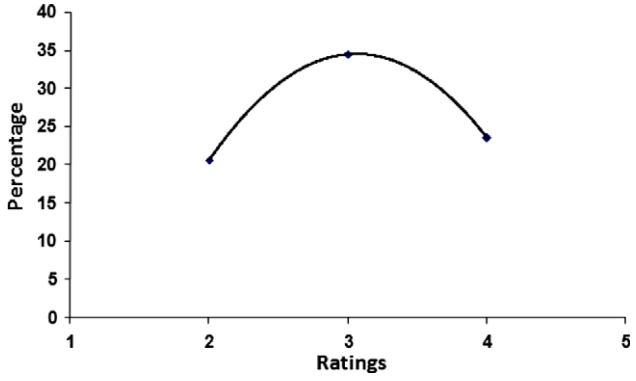

Fig. 12. Corresponding bottom-ranked without extremes in Douban.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

film. The Chinese, on the other hand are demonstrably less likely to give extreme ratings, perhaps because they are more influenced by the consensus and the average sentiment of the reviews already posted or alternatively perhaps because they are not accustom to express extreme emotions.

Overall there is a huge difference between consumer reviews of all the movies in IMDB and Douban. This difference is visible when comparing reviews of all movies, that is, when comparing Figs. 7 and 8. But it is most visible and most pronounced when comparing reviews of the bottom-ranked movies, that is when comparing Figs. 11 and 12. It seems that American reviewers with the most extreme opinions are greatly over-represented relative to the population at large, which distorts American movie ratings relative to the true perceived quality. These results lead us to believe that there is less under-reporting bias among Chinese reviewers, as shown by the bell-shaped curve in Fig. 8. We were concerned that this effect could also be explained not by under-reporting bias, but by profoundly different responses to the movies when seen by Chinese and American audiences, in which case Chinese viewers might have far more average assessments than American audiences, rather than different propensity to review based on the strength of their assessments. To address this concern, we conducted a survey to compare the distribution of assessments of the two populations, to verify whether the difference in reviewing behavior was due to different levels of under-reporting bias or to some other, perhaps as yet unreported, behavioral difference.

5.2. Hypotheses testing

5.2.1. Impact of attitude and social norms on rating behavior

To examine the proposed hypotheses, the first experiment was designed to study how ratings are influenced by general consensus:

$$Rating = f(Average\ rating, Average\ rating\ in\ IMDB\ Bottom\ movies, Average\ rating\ in\ IMDB\ Top\ movies, Average\ rating\ in\ Douban\ Top\ movies,\ t)$$

This is translated into the following empirical model:

$$\begin{aligned} Rating_{jt} = {} & \alpha_1 \cdot AvgRating_{j,t-1} + \alpha_2 \cdot AvgRating_{j,t-1} \\ & \cdot IMDB\_Bottom + \alpha_3 \cdot AvgRating_{j,t-1} \cdot IMDB\_Top \\ & + \alpha_4 \cdot AvgRating_{j,t-1} \cdot Douban\_Top + \alpha_5 \cdot t + \alpha_6 \\ & \cdot Bottom\_Dummy + \alpha_7 \cdot IMDB\_Dummy + \varepsilon_{jt} \end{aligned} \quad (1)$$

In this model, t is the sequence order of each review to control the temporal effect. The first review posted for a movie will have $t = 1$. $AvgRating_{j,t-1}$ denotes the average consumer rating at the time when the $(t-1)$th review was written for movie item j. *IMDB_Bottom* is a dummy variable equal to 1 for the ratings of bottom-ranked movies in IMDB. *IMDB_Top* is also a dummy variable equal to 1 for the ratings of top-ranked movies in IMDB. *Douban_Top* is a dummy variable equal to 1 for the ratings of top-ranked movies in Douban. *Bottom_Dummy* is a dummy variable equal to 1 for movies in the bottom-ranked category. Finally, *IMDB_Dummy* is a dummy variable equal to 1 for movies in the IMDB category.

We estimated Model 1 using robust regression procedure (Chen 2002, Yaffee 2002) to study how ratings are influenced by online review environment variables and presented the results in Table 3.[8] None of the between variable correlations is larger than 80% and the condition index is about 35. A further look at this condition index shows that it is related to the correlation between the main effect and the interaction terms. Hence multicollinearity is not much of concern in our analyses. If we consider a boundary of −1 and 1, our results show that for both cultures, raters have different rating tendency for the top and bottom-ranked movies. Particularly, the parameter estimate for average rating is significantly positive (*AvgRating* = 0.9279 and $p < 0.001$). This means that when Chinese consumers rate a bad product (in this case a movie), if the consensus rating is −1, new consumers will give a rating of −0.9279, a slightly less negative score that is within the boundary of −1 and 1. The same principle applies when Chinese reviewers rate a good movie; if the consensus rating is 1, new reviewers will give a rating of only about 0.24, which is the sum of 0.9279 and −0.6917. This, again, is a less positive score with a tendency to rate towards the center. Overall, the results show that Chinese reviewers tend to leave ratings within the boundary of the general consensus and are not likely to post a rating that is more extreme than the average of what the community has already given. This provides support for the Societal Norms Hypothesis (H3).

Table 3
Average ratings in IMDB and Douban.

Variable	Model
AvgRating	0.9279***
*Avg_Rating * IMDB_Bottom*	0.2631*
*AvgRating * IMDB_Top*	0.1738
*AvgRating * Douban_Top*	−0.6917***
Intercept	3.4503***
Sequence	−0.0033***
IMDB dummy	1.6835***
Bottom dummy	−3.3391***

* $p < .10$.
** $p < .05$.
*** $p < .01$.

However, for the US reviewers, the story is different, especially when they are facing a movie that they perceive to be of low quality. The interaction between *IMDB_Bottom* and average rating is positive (*AvgRating * IMDB_Bottom* = 0.2631 and $p < 0.001$). This indicates that when American consumers rate a bad movie, if the consensus rating is −1, new consumers will give a rating of −1.19 (which is the sum of −0.9279 and −0.2631) a more negative score with tendency to move out of the boundary of −1 and 1. Overall, it seems that US reviewers are not confined to the rating boundary of the community, and they are more willing to post extreme reviews, especially when they are dissatisfied. Perhaps Americans value the need to clearly express their likes and dislikes for movies and are less influenced by the mean, which again gives support for the Freedom of Expression Hypothesis (H2).

5.2.2. Motivation for writing online reviews

For this section, we examine what motivates consumers to post online movie reviews. In our survey, we asked the respondents: (1) If they have been to movie review websites? (2) If they have ever rated a movie in movie review websites? (3) Under what circumstances would they rate a movie online—(a) when they like the movie? (b) When they dislike the movie or (c) when they want to share their opinions with others? (4) How much influence do online movie reviews have on them? (5) How often do they rate movies?

From the survey results collected, we perform a logistic regression to study the motivation for consumers to write online reviews.[9] Our dependent variable is the frequency of rating movies

[8] We used the PROC ROBUSTREG procedure in SAS, which attempts to unweight outlying observations and calculate stable and resistant estimators using robust regression techniques. This solution addresses the non-normality and heteroskedasticity issues created by outliers (Chen 2002, Yaffee 2002).

[9] We also conduct our analysis using ordinary least squares and robust regression procedures. The results were qualitatively similar.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

Table 4
Motivation to post ratings.

Variable	Model
Been_to	1.5663**
Rated	1.1396***
Like	0.6130**
Dislike	0.0604
Share	0.6647**
Influence	0.0331
Intercept	−3.6553***

* $p < .10$.
** $p < .05$.
*** $p < .01$.

by each respondent (Question 5), where 1 means 'never rated' and 5 represents 'often rated'.[10] Our dependent variables capture the answers for Questions 1–4[11]:

$$Frequency = f(Been\_to,\ Rated,\ Like,\ Dislike,\ Share,\ Influence)$$

In Table 4, the variables of "*Been_to*" and "*Rated*" are control variables, while the remaining are our variables of interest. We see that the desire to share opinion with others (*Share* = 0.6647 and $p < 0.05$) dominates all other expressed reasons in motivating a consumer to write a review, giving strong support for the Social Needs Hypothesis (H5), at least among reviewers from collectivist culture. The next factor that motivates consumers to write online movie reviews is the desire to express their liking for the movie (*Like* = 0.6130, $p < 0.05$). Consistent with Hofstede's cultural classifications and with our hypotheses, when Chinese raters like a movie, they are more inclined to post online reviews; hence the much smaller numbers of truly negative online reviews posted on Douban. Finally, individuals from collectivist culture do not appear to write online reviews because of the need for esteem or to influence the views of others, rejecting the Esteem and Recognition Hypothesis (H6). To summarize, consumers from collectivist culture are more likely to speak out when they like a movie instead of when they dislike a movie.

5.2.3. Under-reporting bias

In the previous section, we documented that there are huge cultural influences affecting the behavior of Chinese and American customers when they post online movie reviews. In this section, we study whether such behavioral differences affect the accuracy of the average of reviews as an indicator of the broader group's perception of quality. Given that prior research (Hu et al. 2006) has found the existence of under-reporting bias in United States consumers, causing online reviews to be a biased estimator of books' perceived quality, we study whether this is also true for movies and whether this varies across different cultures. We found that American reviews did indeed exhibit considerable under-reporting bias even for the average-ranked movies, but as expected we found much less under-reporting bias among Chinese reviewers.

As described in Section 4.2, we obtained an additional set of reviews that we were reasonably certain that these reviews did not exhibit under-reporting bias. Having a set of reviews that we were confident did not exhibit under-reporting bias allowed us to determine if Americans posted more extreme reviews due to under-reporting bias or actually held more extreme views than Chinese for some as yet unexplained reason. To obtain the set of reviews that are less likely to be subjected to under-reporting bias, we first try an online survey approach. Survey request was sent to 500 email addresses collected from multiple websites using an email spider program. Due to the low response rate, we were unable to use the results of the survey. As an alternative, we used ratings from MovieLens in place of the survey results to assess the true perceived quality for US movie-goers. Since the aim of users in MovieLens is to improve the quality of their received recommendations, they are more likely to rate all the movies that they have watched. Thus, there is much less reporting bias inside MovieLens. In total, we obtained a dataset that contains 1000,209 movie ratings for 3900 movies contributed voluntarily by reviewers to MovieLens.

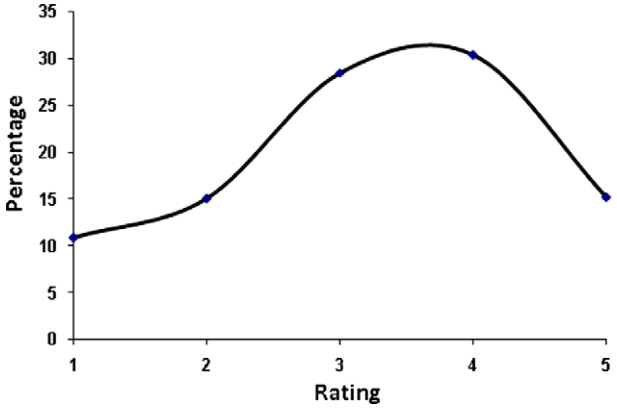

Fig. 13. MovieLens rating distribution with average rating of 3.

Likewise, we focus on average-ranked movies and plot the rating distribution in Fig. 13, which is given by the respondents in MovieLens for 685 movies with average rating of 3. And indeed from Fig. 13, we see that the rating distribution is still unimodal. In contrast, the online review rating distribution we obtained for the average-ranked movies in IMDB on a 5-point scale is W-shaped as shown in Fig. 7. This confirms the existence of under-reporting bias in American movie reviewers, and such reporting bias does indeed cause online reviews to be a biased estimator of a product's true quality as perceived by the broader population of American movie-goers.

Next, we sought to verify whether such under-reporting bias exists in the other population we studied -the Chinese reviewers in Douban. Again, as described in Section 4.2, the existence of under-reporting bias among Chinese reviewers was examined using a set of controlled experiments in which all respondents were asked to report their ratings for several movies that they have viewed. Figs. 14 and 15 show the survey results for the Chinese students when assessing the top and bottom-ranked movies respectively. The survey results revealed patterns similar to the results gathered from the online website, Douban.com. In particular, the results obtained from the online movie reviews in Douban (Fig. 6) and the results from our survey (Fig. 15) follow a similar bell-shaped distribution for the bottom-ranked movies. By comparing the movie ratings in Douban for the top-ranked movies (Fig. 4) with the results from the survey (Fig. 14), we can likewise conclude that the rating distributions are similar across these two channels for bottom-ranked movies as well. Furthermore, the mean difference between online ratings and offline ratings is insignificant ($t = 0.45$ and $p = 0.6569$).[12] Our interpretation for this is that there is far less

[10] Our dependent variable is on the Likert scale of 1–5 where 1 means 'never rated', 2 means 'rated once', 3 means 'seldom', 4 means 'sometimes' and 5 means 'often rated'.

[11] Question 1 is for the first dependent variable "*Been_to*"; Question 2 for the second dependent variable "*Rated*"; Question 3a for the third dependent variable "*Like*"; Question 3b for the fourth dependent variable "*Dislike*"; Question 3c for the fifth dependent variable "*Share*" and Question 4 for the sixth dependent variable "*Influence*".

[12] The *p*-value for the *F*-test of equal variances is 0.7915, therefore we cannot reject the null hypothesis that the underlying variances of the observations are equal.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

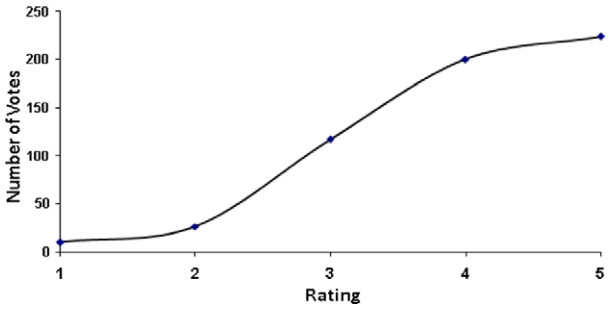

Fig. 14. Chinese students survey data for top-ranked movies.

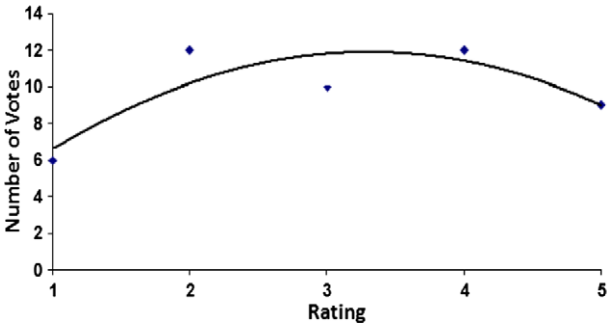

Fig. 15. Chinese students survey for bottom-ranked movies.

under-reporting bias for Chinese online raters, and the overall online Chinese consumer opinion is a well representation of a product's true quality, as measured by the average perception of the broader population of Chinese movie-goers.

We have shown that the extent of under-reporting bias does indeed vary across cultures. In comparison to the results obtained from the United States reviewers, the existence of under-reporting bias is less severe among Chinese movie reviewers. Hence, the average posted online ratings from Douban appear to be similar to those of the "silent consumers" who did not provide their ratings. However, in the case of IMDB, there seems to be far more online postings from movie-goers with the most extreme and indeed the most extremely negative views, and thus the set of posted reviews may not be truly and accurately representative of the "silent consumers" in the United States.

Most of the findings in this study are in accordance with expectations based on Hofstede's cultural dimensions and his characterization of both American and Chinese cultures. As hypothesized, the differences in attitudes towards a movie—that is, differences in underlying assessments of the movie—had very different effects on the behaviors for online reviewing among collectivist and individualist populations. An explanation could lie in the influence of individualism in which individualists perceive that they are relatively free to follow their own wishes and outwardly express them. The fact that the percentages of reviews for the top-ranked and bottom-ranked movies in the US sample are much higher than those of the Chinese sample is consistent with this assertion.

In terms of societal norms, collectivists display a much stronger adherence to the consensus of their communities, including of course the consensus of their online network communities. For the Chinese reviewers, there is concern for reconciliation, harmony, and balance. This may result in vague expression of personal emotions such as likes and dislikes.

To test if the results based on Chinese reviewers hold for other collectivist societies, we replicated our study in Singapore, which has been found to be more collectivist than individualist due to the Confucian heritage of the majority of the population (Hofstede 2001).

6. The Singapore data

In this section, we examine how cultural elements influence the attitudes and intentions in the hybrid culture of Singapore. Although three-fourths of the Singapore population is Chinese, Singaporeans undergo a British system of education, with English being the main medium. Due to their colonial history, Singaporeans have been influenced by Western culture, but Eastern culture and values are also strong. Obedience, harmony and concern for reconciliation are important. Cultural factors have been shown to be mediators of attitude and behavior in Singapore (Tan and Farley 1987).

While Singaporean students are generally more exposed to Western values than their parents were, they still do possess traditional Chinese values as well. This perhaps explains why the results for the top-ranked movies (Fig. 16) in the survey of Singaporean students were similar to those in Douban (Fig. 4). In particular for the bottom-ranked movies in Fig. 17, the experimental results revealed unimodal distribution with mostly moderate reviews, which is similar to the result for the Bottom 100 Movies in Douban.com, as in Fig. 6.

80.5% of our survey participants have been to movie review websites, but only 14.2% have posted ratings. When asked when they would be most likely to provide movie review ratings, 66.5% responded that they would if they liked the movie very much and only 29.1% responded that they would if they were very disappointed. This suggests that Singapore reviewers still exhibit characteristics of their parents' collectivist culture, similar to the behaviors we observed in our Chinese dataset. Once again we observe that reviewers in a collectivist culture are less likely to express their dissatisfaction. Similarly, by examining the data, we conclude that Singaporean reviewers are more reserved about giving the highest ratings even for the top-ranked movies. Once again, we observe support for the Collectivist Societies Hypothesis (H1).

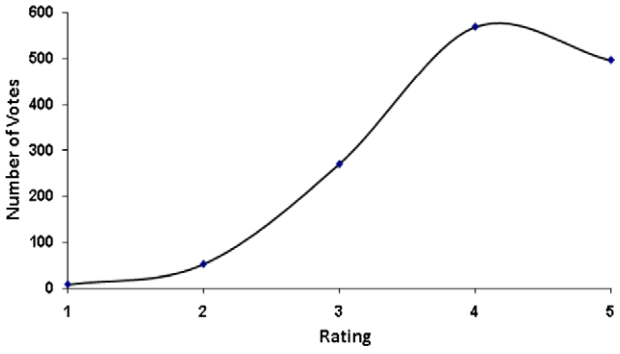

Fig. 16. Singaporean students survey for top-ranked movies.

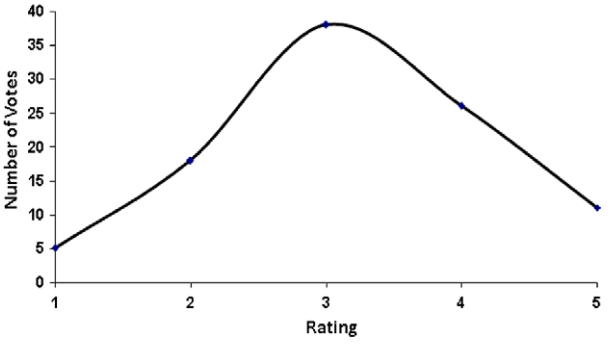

Fig. 17. Singaporean students survey for bottom-ranked movies.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

7. Robustness of the results

In the previous sections, the analyses are based on the online dataset obtained through two rounds of data collection activities. For the first round of data collection, reviews were collected on 1000 movies randomly selected from IMDB and 1000 randomly selected from Douban. We used these datasets to investigate the rating differences across these two websites. However, our conclusions might have been influenced by the differences between the movies selected from these websites. To control the movie title effect, we therefore collected another round of data. For this round of data collection, we first gathered the movie reviews of the top-ranked 100 and bottom-ranked 100 movies in IMDB, then based on these movie titles we collected the related movie reviews from Douban.com. However, one might again argue that the top-ranked and bottom-ranked movies in IMDB may not be top-ranked and bottom-ranked movies in Douban, which might once again have contributed to the observed differences.

Hence, we conduct a robustness check to ascertain if the results are consistent if we were to take the top and bottom-ranked movies in Douban. Based on the same movie titles collected from Douban, we collect the same movies in IMDB and finally the related reviews from both sites. Since Douban does not have a bottom-ranked movie list, we had to scan manually through all the movies in Douban to find the bottom-ranked 100 movies, those with the lowest average rating. Then, from the top-ranked and bottom-ranked movie list in Douban, we extract the corresponding movies in IMDB and plot the rating distribution for these movies as shown in Figs. 18–21. We find that the results were consistent with what we had previously obtained, and that once again extreme ratings are more prevalent among the online reviews written by US movie raters, regardless whether the movies are top-ranked or bottom-ranked at either website.

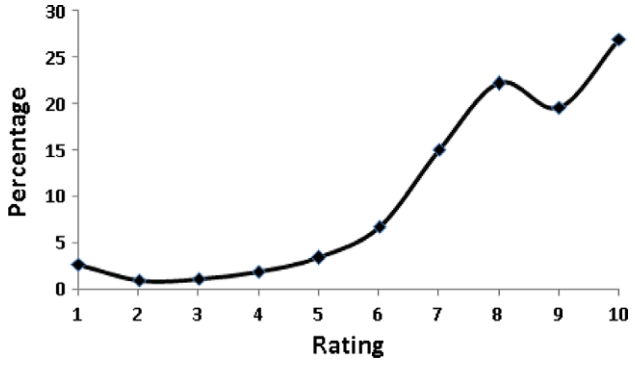

Fig. 18. Corresponding (top-ranked) IMDB movie ratings.

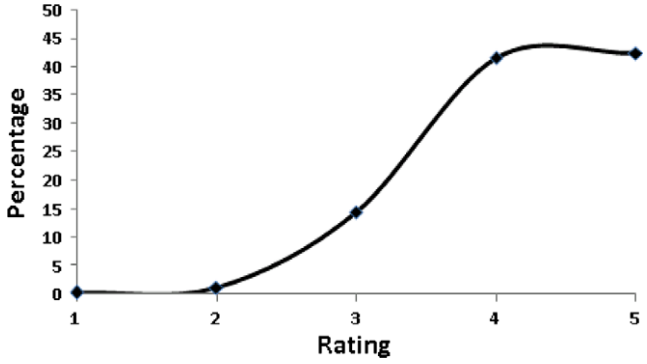

Fig. 19. Douban top-ranked movie ratings.

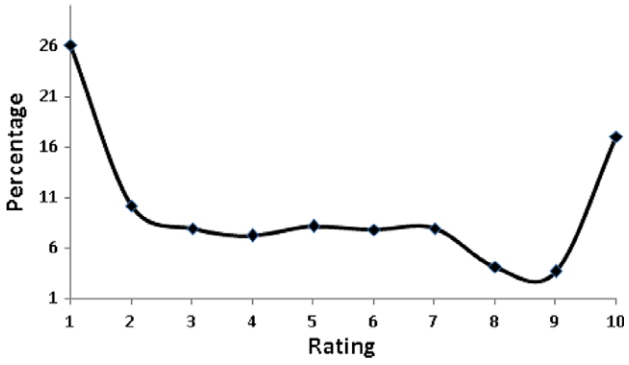

Fig. 20. Corresponding (bottom-ranked) IMDB movie ratings.

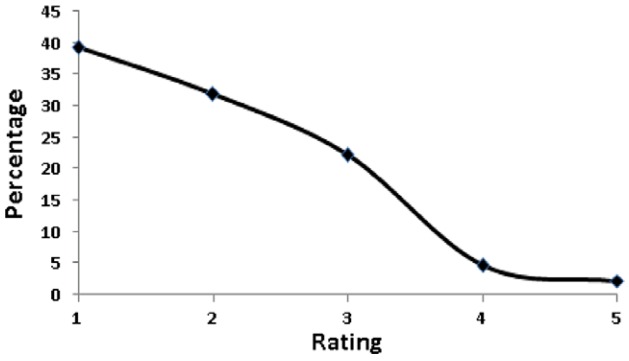

Fig. 21. Douban bottom-ranked movie ratings.

8. Implications for theory and research

This study contributes to our understanding of the role of social norms on individuals' behavior when writing online movie reviews and, we believe, to our understanding of the role of social norms in social networks more generally. Based on the empirical study conducted over three different population groups, from China, the United States, and Singapore, we find that under-reporting bias varies across cultures and cultural differences play a significant role in online reviewing behavior.

The main contributions come from applying Hofstede's (2001) cultural classifications, which predict certain differences in behaviors across cultures, and using these classifications to analyze differences in online behavior in a specific setting. We are not attempting to test Hofstede's theory, and we do not argue that all Chinese are collectivist, or that all Westerners are extreme individualists; we simply analyze millions of movie reviews posted online on websites that cater primarily to Chinese and American reviewers, and look for differences. The differences are indeed consistent with predictions based on Hofstede's theory: Western reviews are much more likely to be extreme, and their distribution tends to become more extreme over time, while Chinese reviews tend to have a more bell-shaped distribution and newer additional posts are much more likely to be closer to the mean rather than more extreme. That is, in Western reviews we observe far more under-reporting among reviewers with average opinions. Such results are further validated by comparisons between the online and offline consumer reviews.

Movies have always been made principally for their home markets, but American movie producers in general hope for more global appeal and more global commercial success. This study indicates that online social reviewing behavior differs greatly from market to market, and might indeed lead a film's distributor to

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001

misjudge the size of a potential market abroad. In particular, a distributor based in one market will know how to interpret early reviews in his home market, but if he applies his home-market experience to interpret the reviews from a foreign market he may be greatly misled. Reviewer behavior at home market that indicates a moderately successful film might be associated with market failure, or with blockbuster success in another market. Thus, an American distributor might over-estimate the market in Singapore or in China, given the greater tolerance of reviewers, or, conversely, a Chinese distributor might under-estimate the market in America given the extreme behavior of some American reviewers. Most importantly, after comparing Figs. 3 and 4, we realize that an American distributor might significantly *under-estimate* the market for a hit American movie in Singapore or China since Fig. 4 does not exhibit the spike that correspond to American reviewers *over-reporting* of reviews of 9 and 10 for top-ranked movies. Likewise, a Chinese movie distributor might note a huge number of extremely negative reviews, an order of magnitude more than he might expect in his home market, and conclude that the launch would be catastrophic. If online behavior is not representative of offline behavior, and if the differences between online and offline behavior vary by nation, then the information in online networks needs to be interpreted carefully before these reviews can be of use to either the community or marketers.

Finally, there are several limitations of this study. First, several parts of the study were performed using students; these portions of the study should be replicated with a non-student population. Second, we were unable to examine the individualists' motivation for posting reviews due to insufficient responses and we feel cultural differences may be at work here. Thus, we hope to gather sufficient responses in future to conduct further analysis on this part. Third, we did not examine the text comments for the posted reviews. It is likely that the extreme opinions of the Chinese are reflected only in the text comments instead of the numerical ratings. Therefore, future research could apply sentiment analysis techniques on the text comments to enable more comprehensive analyses. Also, it might be useful to see how similar results are in retail websites such as Amazon, eBay and others. Finally, further research on the behavior of Americans and of Chinese in online social networks and blogs not associated with commercial purposes would help to strengthen our understanding of cultural differences online.

Acknowledgment

This research was funded in part by the Wharton-SMU Research Center at Singapore Management University.

References

Ajzen, I. From intentions to actions: a theory of planned behavior. In J. Kuhl and J. Beckmann (eds.), *Action Control: From Cognition to Behavior*, Springer-Verlag, New York, NY, 1985.

Ajzen, I. *Attitudes, Personality, and Behavior*. Dorsey Press, Chicago, IL, 1988.

Ajzen, I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 1991, 179–211.

Alexa Internet. Top sites China, 2008. <www.alexa.com/site/ds/top_sites?cc=CN&ts_mode=country&lang=none>.

Anderson, E. W. Customer satisfaction and word of mouth. *Journal of Service Research*, 1, 1, 1998, 5–17.

Bickart, B., and Schindler, R. M. Internet forums as influential sources of consumer information. *Journal of Marketing*, 15, 3, 2001, 31–40.

Bond, R., and Smith, P. B. Cross-cultural social and organizational psychology. *Annual Review of Psychology*, 47, 1996, 205–235.

Chatterjee, P. Online reviews: do consumers use them? *Advances in Consumer Research*, 28, 1, 2001, 129–133.

Chen, C. Robust regression and outlier detection with the ROBUSTREG procedure. In *Proceedings of the Twenty-seventh Annual SAS Users Group International Conference*, SAS Institute Inc., Cary, NC, 2002.

Chevalier, J., and Mayzlin, D. The effect of word of mouth on sales: online book reviews. *Journal of Marketing Research*, 43, 3, 2006, 345–354.

Clemons, E. K. How information changes consumer behavior and how consumer behavior determines corporate strategy. *Journal Management Information Systems*, 25, 2, 2008, 13–40.

Clemons, E. K., and Gao, G. Consumer informedness and diverse consumer purchasing behaviors: traditional mass-market, trading down, and trading out into the long tail. *Electronic Commerce Research and Applications*, 7, 1, 2008, 3–17.

Clemons, E. K., Barnett, S., and Appadurai, A. The future of advertising and the value of social network websites: some preliminary examinations. In C. Dellarocas and F. Dignum (eds.), *Proceedings of the 9th International Conference on Electronic Commerce, Minneapolis, MN, August 2007*, ACM Press, New York, NY, 2007, 267–276.

Dellarocas, C. The digitization of word-of-mouth: promise and challenges of online feedback mechanisms. *Management Science*, 49, 10, 2003, 1407–1424.

Ess, C., and Sudweeks, F. Culture and computer-mediated communication: toward new understandings. *Journal of Computer-Mediated Communication*, 11, 1, 2005. <http://www.jcmc.indiana.edu/vol11/issue1/ess.html>.

Fishbein, M., and Ajzen, I. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Addison-Wesley, Reading, MA, 1975.

Hofstede, G. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Sage Publications, Thousand Oaks, CA, 1980.

Hofstede, G. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*, second ed. Sage Publications, Thousand Oaks CA, 2001.

Hofstede, G., and Bond, M. H. The Confucius connection: from cultural roots to economic growth. *Organizational Dynamics*, 16, 4, 1988, 4–21.

Hu, N., Pavlou, P. A., and Zhang, J. Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication. In J. Chuang and D. M. Pennock (eds.), *Proceedings of the 7th ACM Conference on Electronic Commerce, Ann Arbor, Michigan, June 11–15, 2006*, ACM Press, New York, NY, 2006, 324–330.

Lanier, J. *You are not a Gadget: A Manifesto*. Alfred A. Knopf, New York, NY, 2010.

Li, C., and Bernoff, J. *Groundswell: Winning in a World Transformed by Social Technologies*. Harvard Business Press, Boston, MA, 2008.

Madden, T. J., Ellen, P. S., and Ajzen, I. A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, 18, 1, 1992, 3–9.

Maslow, A. H. A theory of human motivation. *Psychological Review*, 50, 4, 1943, 370–396.

Nakata, C., and Sivakumar, K. Instituting the marketing concept in a multinational setting: the role of national culture. *Journal of the Academy of Marketing Science*, 29, 3, 2001, 255–275.

Pavlou, P. A. What drives electronic commerce? A theory of planned behavior perspective. In *Best Paper Proceedings of the Academy of Management Conference*, Denver, CO, August 2002.

Pavlou, P. A., and Lin, C. What drives electronic commerce across cultures? A cross-cultural empirical investigation of the theory of planned behavior. *Journal of Electronic Commerce Research*, 3, 4, 2002, 240–253.

Poulton, E. C. *Bias in Quantifying Judgments*. Erlbaum, Hillsdale, NJ, 1989.

Senecal, S., and Nantel, J. The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, 80, 2004, 159–169.

Sondergaard, M. Research note: Hofstede's consequences: a study of reviews, citations and replications. *Organization Studies*, 15, 3, 1994, 447–456.

Tan, C. T., and Farley, J. U. The impact of cultural patterns on cognition and intention in Singapore. *Journal of Consumer Research*, 13, 4, 1987, 540–544.

Taylor, S., and Todd, P. A. Understanding information technology usage: a test of competing models. *Information Systems Research*, 6, 3, 1995, 144–176.

Tourangeau, R., Rips, L. J., and Raskinski, K. *The Psychology of Survey Response*. Cambridge University Press, 2000.

Wu, F., and Huberman, B. Novelty and collective attention. *Proceedings of the National Academy Science*, 104, 45, 2007, 17599–17601.

Yaffee, R. A. *Robust Regression Analysis: Some Popular Statistical Package Options*. Statistics Department, New York University, New York, NY, 2002. <http://www.nyu.edu/its/statistics/Docs/RobustReg2.pdf>.

Please cite this article in press as: Koh, N. S., et al. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electron. Comm. Res. Appl.* (2010), doi:10.1016/j.elerap.2010.04.001