International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

A Source You Prefer, or Majority? Investigating User Responses to Conflicting Opinions in Multi-Platform Restaurant-Review Lists

Tzu-Hao Lin, Yen-Yun Liu, Hong-Han Shuai, Fang-Hsin Hsu & Yung-Ju Chang

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2022.2090613

A Source You Prefer, or Majority? Investigating User Responses to Conflicting Opinions in Multi-Platform Restaurant-Review Lists

Tzu-Hao Lin[a], Yen-Yun Liu[b], Hong-Han Shuai[c], Fang-Hsin Hsu[d], and Yung-Ju Chang[a]

[a]Department of Computer Science, National Yang Ming Chiao Tung University, Hsinchu, Taiwan (R.O.C.); [b]Department of Social and Behavioral Sciences, Utrecht University, Utrecht, The Netherlands; [c]Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan (R.O.C.); [d]Department of Network Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan (R.O.C.)

ABSTRACT
Customers learn about restaurants in various ways, and integrating this disparate information could give them access to a greater diversity of perspectives. Conflicting opinions between restaurant-review platforms are inevitable. However, such conflicts' influences on users' perceptions remain unclear, especially when the opinion of a user's preferred platform conflicts with the majority of others. This study's experiment with a sample of 304 users found that, when such situations occurred, the preferred platform's influence differed depending on whether the user was shown a sequence of whole-platform aggregations vs. a sequence of individual reviews drawn from multiple platforms. That is, the participants accepted the majority view most of the time, but when looking at aggregated lists, if their preferred platform expressed a minority positive opinion based on a high quantity of reviews, that minority opinion could prevail over the majority one. Between-platform conflicts were also found to have a greater impact on user reactions than within-platform ones did.

1. Introduction

As a type of electronic word of mouth (eWOM) communication, online consumer reviews have become crucial to evaluations of the quality and performance of certain products and services (Chen & Xie, 2008; Filieri, 2015). However, online reviews have been found susceptible to active manipulation (Zannettou et al., 2019). Even where such manipulation does not occur, consumers may be biased toward the preferences of a particular source, including their media preferences (Iyengar & Hahn, 2009). To mitigate the influence of such bias, researchers have studied the potential benefits of exposing people to diverse opinions and perspectives from various information sources (e.g., Park et al., 2009; Ribeiro et al., 2018), on the basis that information from multiple sources is more likely to be based on independent pools of knowledge, making it worthier of consideration (Harkins & Petty, 1987). However, while integrating diverse sources is likely to expose users to diverse perspectives, such diversity naturally tends to be marked by conflicting opinions. In a context such as restaurant reviews, seeing conflicting opinions may lead to confusion, and/or cause users to conclude that such information is not helpful to decision-making (Baek et al., 2012; Qiu et al., 2012).

While prior research on users' perceptions of conflicting reviews mainly focused on "within-source" reviews, i.e., reviews from the same platform, how users perceive conflicting opinions among *different* media/information sources in mobile local searches has not yet been explored (Fan & Chen, 2014). In particular, while people's biases and preferences regarding information sources have been found to affect their acceptance of others' opinions when they are exposed to news from a variety of such sources, it is unclear whether such biases and preferences also affect geo-linked opinions, such as reviews of specific places. And, when people are aware of the opinion of their preferred platform conflicts with the majority of other reviewers' opinions, which side will they take? To fill this research gap, the current study aimed to capture users' perceptions of conflicts in multi-source review lists, with special attention to the influence of review conflicts on their (1) perceptions of these lists' helpfulness and (2) eventual acceptance of one conflicting opinion over another.

We chose restaurant reviews as our research focus, as they form a frequently sought category of information in daily life (Teevan et al., 2011; Whitla, 2009; Zhang et al., 2010). Additionally, as there are two common types of reviews—aggregations of average ratings, and sequences of individual reviews—we investigated users' perceptions of conflicts in both these types of multi-source reviews. Our first research question is as follows:

> RQ1: How does review conflict in a multi-source restaurant-review list influence users' a) perceptions of that list's helpfulness and b) intention to visit the reviewed restaurant?

CONTACT Yung-Ju Chang armuro@cs.nctu.edu.tw National Yang Ming Chiao Tung University, 1001 University Rd, Hsinchu, 30010, Taiwan (R.O.C.)

Based on our findings regarding such influence, we will then explore how users come to accept one conflicting opinion over another. Hence:

> RQ2-1: When the opinion of a restaurant expressed by a user's preferred review platform conflicts with the majority opinion expressed on other review platforms, a) how does such conflict impact his/her perceptions of that list's helpfulness, and b) which side's opinion does s/he tend to adopt?

Ratings lists are based on aggregations of average ratings, and generally contain additional quantitative information. Because prior research has shown that the quantity of reviews influences readers' perceptions of how persuasive certain reviews are (De Pelsmacker & Janssens, 2007), our experimental conditions also manipulated the quantity of reviews. Thus,

> RQ2-2: When the opinion of a restaurant expressed by a user's preferred review platform conflicts with the majority opinion expressed on other review platforms, does the quantity of reviews from the preferred source have an impact on users' a) perceptions of that list's helpfulness and b) intentions to visit the reviewed restaurant?

Finally, as previous research focused narrowly on differences of opinion among reviews from the same platform, we investigated whether within-platform conflict had the same influence as between-platform conflict did. Therefore,

> RQ3: How do conflicts of opinion between reviews posted on the same platform influence a) users' perceptions of the helpfulness of the list in which such conflicts appear, and b) their intention to visit the reviewed restaurant?

Based on 304 individuals' responses to our online questionnaire, we are able to report several novel findings. First, we found that when the opinion expressed by a user's preferred review platform conflicted with the opinions expressed by other such platforms, the latter had a greater impact on their visit intentions. Second, we found that in such a situation, the influence of a user's preferred review platform varied in strength according to whether the reviews were of aggregations of ratings or sequences of discrete reviews. In the case of the former review type, the influence of preferred-platform minority opinion could be moderated by the quantity of reviews: i.e., a small number of reviews on the respondent's preferred platform did not affect the list's perceived helpfulness, but a large number of positive opinions significantly increased the participants' restaurant-visit intentions. Lastly, reviews' perceived helpfulness arising from opinion conflicts were mainly associated with between-platform rather than within-platform ones. This study may be considered useful to the development of multi-platform reviews by third-party e-retailers such as Tripadvisor.com and Yelp.com.

2. Related work

People tend to process online reviews heuristically (TODOROV et al., 2002), an approach that minimizes the cognitive effort required to make a decision (Chen & Chaiken, 1999; Sparks & Browning, 2011). Characteristics of reviews that act as heuristic cues usually include their sources, quantity, and consistency. It has been suggested, in the case of experience products, that source credibility and specific rating valence are both important to people's evaluations of whether reviews are helpful (Baek et al., 2012; Filieri & McLeay, 2014).

Although people often use heuristic cues to process information, it is suggested that they would think more deeply when there is a conflict between reviews (Ruiz-Mafe et al., 2018). Additionally, because consumers' online reviews often include ratings with open-ended comments (Park & Kim, 2008), the quality of the arguments presented in such comments may also play a role in the perceived usefulness of a product or service. According to Willemsen et al.'s (2011) content analysis of reviews of experience and search products, the density and diversity of argumentation both served as significant predictors of the perceived usefulness of reviews. More specifically, however, consumers found reviews more useful when they were more elaborate and contained both pros and cons of a product. And another study, by Cheung et al. (2012) suggested that quality of argumentation acted as the primary factor of whether a source is believable.

The next five subsections discuss the constructs most relevant to multi-platform review lists, including source credibility, variation in users' preferences, review valence, review consensus, and review quantity.

2.1. Source credibility

Source credibility refers to a person's subjective evaluation of the believability of informational sources (Rieh, 2015; Wilson & Sherrell, 1993). It is closely linked to changes in consumers' attitudes (Wilson & Sherrell, 1993) and their perceptions of whether certain information is useful (Rieh & Danielson, 2007; Sussman & Siegal, 2003). It has also been widely suggested that individuals' past judgments about the platforms on which reviews appear are often used as proxies for the source credibility of the reviews themselves (Brown et al., 2007; Dou et al., 2012; Rieh, 2015; Schindler & Bickart, 2005; Senecal & Nantel, 2004; Slater & Rouner, 1996; Wathen & Burkell, 2002; Xue & Phelps, 2004). Source-credibility assessments can also directly affect reviews' perceived helpfulness (Baek et al., 2012; Cheng & Ho, 2015; Choi & Ok, 2011; Filieri et al., 2018; Li et al., 2013; Sussman & Siegal, 2003) and consumers' purchase intentions ((Jimmy) Xie et al., 2011; Lee, 2009; Zhang et al., 2014). Sources with higher perceived credibility are generally also perceived as more persuasive (Pornpitakpan, 2004; Roy Dholakia & Sternthal, 1977), and they are more likely to be chosen as people's preferred source of online opinions (Dabholkar, 2006).

2.2. Variation in individuals' dining preferences

When reading online reviews of restaurants, individuals make decisions based on various source characteristics and dining scenarios (Birnbaum & Stegner, 1979; Chang, 2011; Rieh, 2015). Prior research has also shown that individuals' dining motivations affect restaurant-visit intentions. While some consumers may choose to dine out purely for the taste of food (i.e., intrinsic motivation), others may go to a restaurant for extrinsic reasons, such as being required to attend a business meeting there (Noone & Mattila, 2010; Ponnam & Balaji, 2014).

In addition, consumers' dining motivations can also have an impact on how they perceive reviews' rating consensus and review quantity. According to Lu et al. (2020), consumers who exhibited extrinsic dining motivations were more likely than their intrinsically motivated counterparts to prefer restaurants with reviews that were consistent. Interestingly, the same study also found that a large number of reviews could mitigate this negative impact on the perception of low-consensus reviews; but this effect was found only among intrinsically motivated diners. It has also been suggested that people generally prefer reviews by their peers, perceiving them as more trustworthy and helpful than those written by marketers, online editors, or experts (Bickart & Schindler, 2001; Huang & Chen, 2006; Li et al., 2013; Smith et al., 2005; Sparks et al., 2013).

2.3. Review valence

It is common to see reviews with different *valences*, i.e., positive or negative, within a product-review list. Prior research has found that positive and negative valence differentially affect people's attitudes and expectations about a product (Liu, 2006). According to Kanouse and Hanson (1987), people tend to display negativity bias. That is, due to their tendency to try to minimize risk and uncertainties, they tend to regard negative reviews as more powerful, persuasive, and helpful in decision-making. This effect is especially pronounced in the case of experience products such as restaurants (Racherla & Friske, 2012), along with skincare products (Hao et al., 2010; Willemsen et al., 2011) and online video games (Yang & Mai, 2010).

It has also been found that consumers tend to use review extremity and valence to assess the trustworthiness of a product (Filieri, 2015). However, Filieri et al. (2019) suggested that extremely negative reviews are perceived as helpful only when they are in-depth and easy to read, and when the reviewer is either an acknowledged expert or otherwise identifiable on some level (e.g., his/her geographic location is disclosed). Products' quality may also play a moderating role in perceptions of the helpfulness of extremely negative ratings. In a recent study by Filieri et al. (2019) of accommodation reviews posted online, it was reported that extremely negative reviews were more likely to be perceived as helpful when a hotel had a certificate of excellence, and/or when its average rating score was high.

Turning now to the influence of positive reviews, studies of music, movies, video games (Pan & Zhang, 2011), hotels (Carlson et al., 2011), and restaurants (Pentina et al., 2018) have suggested that people perceive such reviews as more credible and helpful than negative ones. It has also been suggested that a review set containing mostly positive reviews inspired more positive impressions of its reviews, and increased consumers' purchase intentions more, than a majority-negative review set did (Purnawirawan et al., 2014).

2.4. Review consensus

Review consensus, i.e., the level of agreement among reviews written on the same subject by different contributors, can also affect users' perceptions of all such reviews (Doh & Hwang, 2009; Lee & Cranage, 2014). Cheung et al. (2009) reported that people perceived a review to be more credible if it was consistent with other reviews they had read, but Doh and Hwang (2009) found that inclusion of one negative review within an otherwise positive set of reviews could lead the entire set to be perceived as more credible. According to Quaschning et al. (2015), on the other hand, people perceived a review as more helpful when the rating it gave was consistent with those of other reviews of the same product. And conversely, Baek et al. (2012) and Qiu et al. (2012) found that when a list's average rating and the rating given by an individual were inconsistent, it tended to reduce the perceived credibility and helpfulness of the extremely positive or negative reviews in the list. It has also been suggested that when a set of reviews exhibits high consensus—i.e., most of its reviews are either positive or negative—it is perceived as more useful than review sets with a more centralized/balanced distribution (Purnawirawan et al., 2012).

Cultural values may also play a role in how people perceive review variance. Wu et al. (2021) has found a moderating effect between indulgent and restrained cultures. It is suggested that indulgent consumers are more willing to take risks, as they tend to adopt the information with high variance reviews, while restrained consumers are more likely to adopt as the reviews are more consistent (i.e., low variance). Yet, in spite of the abundance of robust research on the influences of review consistency within various domains, there appear to have been no prior examinations of consensus among reviews across multiple different platforms.

2.5. Review quantity

In prior eWOM studies, information quantity has been found to serve as a popularity signal that affects consumers' perceptions of a product and intentions to purchase it (Filieri et al., 2021; Park & Lee, 2008). Review quantity, sometimes termed "volume" (e.g., numbers of likes, numbers of reviews per restaurant), is generally linked to the number of consumers who have purchased a product, which in turn can guide consumers' heuristic evaluations of the popularity of a product; and as such, a large quantity of reviews tends to cause a rating to be perceived as more credible and reliable (Flanagin & Metzger, 2013; Zhang et al., 2010). A rating based on a large number of reviews can also strengthen consumers' positive or negative perceptions of a product, to a greater extent than a low number of reviews can (Khare et al., 2011). Researchers have also found that sheer quantity of reviews can positively impact the popularity of a product (Liu & Park, 2015), as well as consumers' purchasing intentions (Filieri et al., 2018; Lee, 2009; Park et al., 2007; Zhang

Table 1. Review platforms used in this study.

Platform	Description
Google Maps	A map-based wayfinding system incorporating business reviews
Facebook	A social-media platform for friends
Locals	A fictive platform used by "local people" for reviewing local businesses
News media	An aggregation of real restaurant reviews published in newspapers, magazines, and news blogs

et al., 2010; Zhang et al., 2014). Given that prior studies have focused exclusively on reviews posted on single platforms, the current study investigated rating lists aggregated from multiple platforms.

3. Methods

Guided by our research questions, we built a website that collected participants' responses to four review platforms (Table 1). Three of them were real, whereas the platform called Locals was imaginary. Inspired by the concept of crowdsourcing and by some of the specific features of LocalWiki[1] and Local Guide,[2] Locals simulated people sharing their experiences of local restaurants. Our participants were told that the Locals platform included reviews only from people who lived in the areas where the reviewed restaurants were located, as opposed to tourists or other visitors. We included five dining scenarios—*eating alone*, *friend gathering*, *dating*, *business dinner*, and *travel abroad*—on the grounds that participants might already have their own preferred review platforms for different types of dining occasion.

3.1. The two general types of multi-source reviews

When seeking online information about a restaurant, web users typically find two types of multi-source reviews: (1) an aggregated/summarized rating, consisting of the average of all available individual ratings; and (2) individual reviews, each containing its own rating, often in association with comments. Each of these types is dealt with in detail below.

3.1.1. Multi-source summarized rating lists

Our experiment's *summarized rating lists* (hereafter referred to as "SUMs") provided average restaurant ratings drawn from our four platforms (Figure 1), each of which included both an average rating from all contributors to that platform, and a number indicating a *quantity* specific to that platform: i.e., the number of people who had written reviews of the target restaurant on Google Maps and Locals; the number of friends on Facebook who "Like" or have checked in at it; and the number of news-media stories that have been published about it.

These quantities were grouped into three levels: with *high* meaning that the quantity was higher than a certain threshold that led most participants to perceive the average rating as sufficiently reliable; *low*, meaning that the quantity was so small that most participants would not perceive the average rating as credible enough to act upon it; and *mid*, representing all quantities not fitting into either of the other two categories. We applied source-specific approaches to establishing these quantities. For Google Maps and Locals, this was done by picking quantity data from Google Maps reviews of more than 10,000 restaurants in Los Angeles, Chicago, and New York, but—with reference to prior research findings (Qiu et al., 2012) that an average rating given by more than 96 people is sufficiently representative and credible, we set 96 as the upper threshold, and, following Yang et al. (2016), 10 as the lower threshold. As for Facebook and news media, we could not identify any literature on how many "Likes" or check-ins to a restaurant or how many news reports about it would make people perceive a rating as sufficiently credible. However, considering that users could have different personal thresholds for credible numbers of such metrics on these platforms, we asked our participants to specify their own high and low thresholds (ranging from 1 to 100) for each of them, as part of a pre-study questionnaire. Then, for each user, the SUM generated by the system began by randomly determining a quantity as high, mid, or low, and then—according to the upper and lower thresholds for that quantity category—randomly generated a number. For example, if the system determined that the quantity for the news-media platform in a given SUM should be mid, and a participant reported his/her personal low and high thresholds for credible numbers of news-media reviews to be 3 and 20, respectively, the system generated a number between 3 and 19 for news media for that SUM.

3.1.2. Multi-source individual review lists

Our experiment's *individual review lists* ("INDs") comprised reviews apparently written by individual visitors to a given restaurant, and were ascribed to all four of our review platforms (Figure 1). Each review included a star rating and a comment. In reality, not all platforms allow users to apply a stars-out-of-five rating plus textual comments to their evaluations of restaurants. However, based on an assumption that future INDs could easily convert users' reviews into comparable units such as stars-out-of-five ratings, we used the same general rating system for all platforms to make them more comparable for research purposes. Additionally, to help us identify the influence of within-platform conflicts, it should be noted that each IND we generated presented five reviews, but that one of the source platforms in it appeared twice, while the other three appeared only once each.

Based on our study's research purpose, we tried to control for the influence of textual comments. All sample comments were downloaded from Google Maps, and were associated with the original ratings given. Then, based on the rating generated by our system, a comment associated with that same rating in the real world was randomly paired with it. Two prior studies (Zhang et al., 2010) found that

Figure 1. Examples of SUM (left) and IND (right).

Table 2. PN-Combinations of SUM and IND.

Summarized rating list			Individual review list		
Consensus ratio	PN ratio	Consensus level	Consensus ratio	PN ratio	Consensus level
All the same	4P0N	4	All the same	5P0N	5
	0P4N	−4		0P5N	−5
3:1	3P1N	2	4:1	4P1N	3
	1P3N	−2		1P4N	−3
1:1	2P2N	0	1:1	3P2N	1
				2P3N	−1

comments about food quality and taste were the most valued by readers of restaurant reviews, though comment length has also been found to affect reviews' perceived helpfulness (Mudambi & Schuff, 2010). As such, our INDs excluded comments (1) longer than two lines, (2) containing descriptions other than of the food, and (3) whose valences appeared inconsistent with their original numerical ratings.

3.1.3. Positive-and-negative combinations

To obtain data on how diverse combinations of review valences and sources might differentially affect users' perceptions, we generated all possible combinations of positive and negative reviews for each list, which we termed *PN-Combinations* (see Table 2). That is, each participant was shown a series of review/rating lists, within each of which they saw a randomly assigned PN-Combination. Within each such combination, a level of *consensus* could be readily discerned, based on the absolute value of the difference between the list's positive and negative ratings. For example, a SUM with a PN-Combination of 4P0N, i.e., four positive and no negative reviews, is more consistent than one classified as 3P1N, i.e., three positive reviews and one negative one. The former's consensus value is 4, while the latter's is 2: i.e., 3+(−1). That is, the greater the consensus value, the more consistent the ratings in the list were, and a 2P2N list—consisting of two positive and two negative reviews—would have a consensus value of 0 (i.e., 2+(−2)). Within-platform conflict as reflected in INDs, on the other hand, consisted of two reviews in the same list being from the same platform, but having different valences (e.g., one positive and one negative, both from Facebook). We balanced the number of PN-Combinations for each source. In line with prior work (Carbon & Stanford, 2014; Park & Nicolau, 2015; Qiu et al., 2012), we defined any rating of four or five stars as positive, and any of three stars or lower as negative, for both SUM and IND (see Table 3).

3.2. Study procedure

On entering the experiment website, each participant filled out a questionnaire about his/her basic demographic information: i.e., gender, age, education, occupation, and annual income. All were also asked several pre-survey questions covering how frequently they used each of the four focal platforms for learning about restaurants, as well as for their

Table 3. Valences of rating numbers for SUM and IND.

Valence of rating numbers	Summarized rating list	Individual review list
Positive	4.0–5.0	4, 5
Negative	1.0–3.9	1, 2, 3

personal high and low thresholds for credible quantities of Facebook and news-media reviews, as explained above.

3.2.1. Stage 1: Reporting platform preferences

In the first stage of the experiment, the participants were given five dining purposes in a randomly assigned order, told to imagine that they were seeking restaurant information for each such purpose, and asked to choose their preferred platform for obtaining such information.

3.2.2. Stage 2: Assessing lists

In the second stage, the participants were asked to complete two series of questions. In the first (hereafter, "stage 2-1"), there were 25 rounds, each containing one SUM; and in the second ("stage 2-2"), there were 60 rounds, each containing one IND. In both stages 2-1 and 2-2, the rounds were evenly distributed across the five dining purposes.

The list presented in each round had a specific PN-Combination that was pre-determined and randomly assigned to each participant. The order of the platforms was also randomly determined, to avoid ordering effects. We asked the participants to imagine that they saw the rating/review list when searching for restaurants within the assigned dining scenario. Next, they were instructed to answer a yes/no question as to whether each rating/review in that list was "influential" on their judgment about the restaurant, but told that they must not answer this question if they felt neutral about the rating/review's influence. Then, they reported how helpful each four- or five-item rating/review list was, and whether they would add the restaurant to their "want-to-go" list (as a proxy for their visit intention).

3.2.3. Attention-checking questions

At each stage of the questionnaire, two additional attention-checking questions were added to identify potential 'straightlining' behavior (C. Zhang & Conrad, 2014). The data generated by those participants who were thus identified as engaging in such behavior were removed from further analysis.

3.3. Participants

We recruited participants via Amazon Mechanical Turk (MTurk), a well-respected platform for running virtual experiments (Horton et al., 2011; Paolacci et al., 2010). To ensure high-quality responses, participation was limited to MTurk members who had performed at least 1000 prior tasks at an approval rate above 98%. Some additional eligibility criteria, including having experience of using Google Maps and Facebook to learn about restaurants, were also listed on the recruitment page.

Initially, we recruited 340 participants, but not all of their responses could be treated as valid, due to evidence of possible "straightlining" in 36 cases (C. Zhang & Conrad, 2014). To maintain our data quality, we excluded those 36 individuals' responses, leaving a final sample of 304 participants aged 21 to 60 ($M = 34$), of whom 67.7% were male and 63% had a bachelor's degree or above. Their top three job categories were Science & Technology (33.5%), Business Management & Marketing (17.1%), and Finance (7.2%); 1.9% were students. Around 45% had annual incomes under US$30,000, and 18%, above US$60,000.

3.4. Data preprocessing and analysis

The final dataset included 7600 stage 2-1 and 18,240 stage 2-2 responses. Because each participant had a number of repeated observations that differed from each other, we built mixed-effects logistic-regression models in which participant ID was included as a random effect to account for individual differences. Perceived helpfulness and visit intention were included as the two dependent variables (DVs), and the two predictors (IVs) were the fixed effects of (1) consensus level and (2) the valence of majority opinion. We also included, as a fixed effect, a binary variable representing whether the individual participant's preferred review platform was in the minority of opinion or not.

For examining RQ2-2 specifically, we additionally added another fixed effect, the categorical variable *quantity of the preferred source* (low, mid, high), and tested for an interaction effect between it and the binary variable. The purpose of this was to facilitate our exploration of whether the influence of users' review-platform preferences on their restaurant and list perceptions changed based on review-quantity information.

Lastly, to answer RQ3, regarding the effect of within-platform conflict, our model included categorical variables representing overall consensus (high/low) and whether within-platform conflict was shown. All the IVs mentioned above were treated as fixed effects in the models.

4. Results

4.1. Variation in platform preferences across dining purposes

Overall, Google Maps was our participants' most-preferred platform for restaurant-review information. However, individual participants' platform preferences varied depending on their dining purposes (Figure 2). Although the Locals platform was imaginary, the participants seemed able to make sense of the information it provided. Overall, it was the second most-preferred platform, though it was less popular when the user's dining purpose was dating or a gathering of friends. Facebook was preferred more often than Locals for dating, and was the most-preferred platform when the dining purpose was a gathering of friends. News

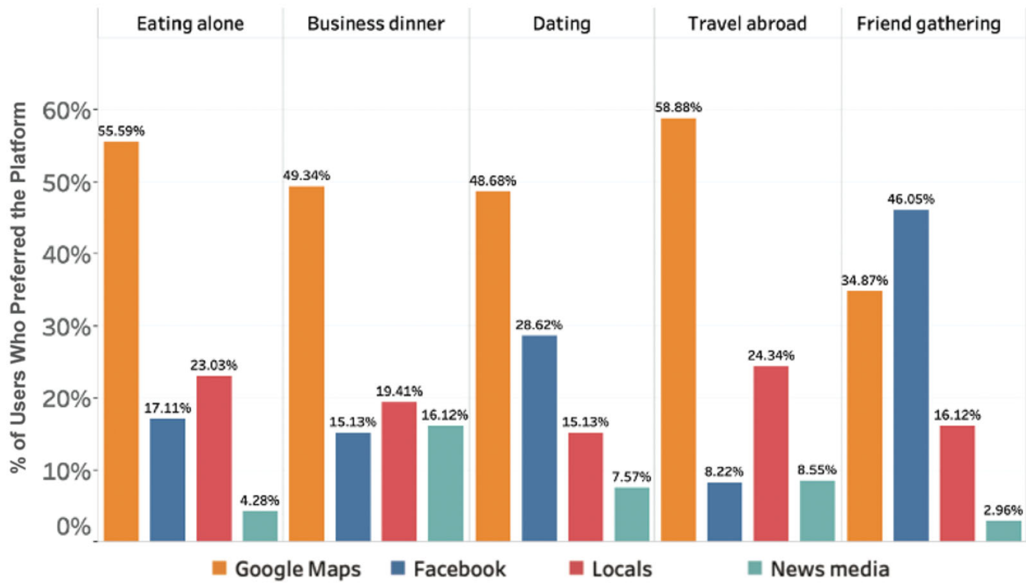

Figure 2. Variation in platform preferences by dining purpose.

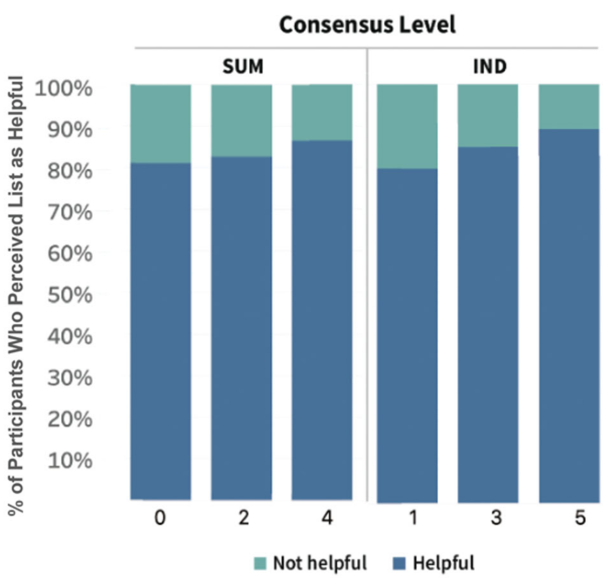

Figure 3. Consensus levels' effects on perceived helpfulness.

media was the least-preferred platform for all dining purposes except business dinners. This was not unexpected, as prior research has shown that when seeking information about experience products such as restaurants, people consider reviews from their fellow consumers to be more credible than those from commercial entities, and may even regard restaurant evaluations from news media as advertisements (Bae & Lee, 2011; Goldsmith & Horowitz, 2006). With regard to the business-dinner scenario in particular, it is possible that the participants felt news media were more likely than other platforms to recommend restaurants with a formal atmosphere.

4.1.1. Level of consensus

The participants reported our multi-source review lists as helpful 80% of the time. However, as shown in Figure 3, the higher its consensus level was, the more helpful the participants perceived a given rating/review list to be, regardless of whether it was in the SUM or IND format. Regression results showed a positive main effect of consensus level on perceived helpfulness (SUM: $Z = 4.263$, $p < .001$; IND: $Z = 17.32$, $p < .001$).

Interestingly, at the same consensus level (e.g., 2 vs. −2 in Figure 4, left), our participants perceived positive-skewing lists as more helpful than negative-skewing ones; and this effect was highly significant (SUM: $Z = 10.471$, $p < .001$; IND: $Z = 15.79$, $p < .001$). This echoed prior work on single-platform reviews (Carlson et al., 2011; Pan & Zhang, 2011; Pentina et al., 2018).

4.1.2. Opinion conflicts between the preferred platform and others

When seeking to answer RQ2, regarding how perceptions of a list's helpfulness are impacted by one's preferred review platform espousing a minority opinion, we only considered circumstances in which such minority opinion was either positive or negative (i.e., 1 P or 1 N in a SUM list, and 2 P or 2 N in an IND list). That is, in an IND list containing two reviews from the participant's preferred platform, one of them positive and one of them negative, the data regarding that list was excluded from our opinion-conflict analysis. As shown in Figure 5, the perceived helpfulness of lists containing such conflicts was quite high. Nevertheless, we still observed a small effect of conflict. That is, participants tended to perceive the helpfulness of a given list as slightly lower when their preferred platform espoused a minority opinion (SUM: 81.19%; IND: 79.74%) than when it did not (SUM: 83.64%; IND: 82.62%). However, this difference was only statistically significant for IND lists ($Z = -4.94$, $p < .001$), not SUM ones ($Z = -1.745$, $p = .0809$).

4.1.3. Quantity of ratings on the preferred platform

As illustrated in Figure 6, when we looked at the influence of the quantity of ratings on the preferred platform

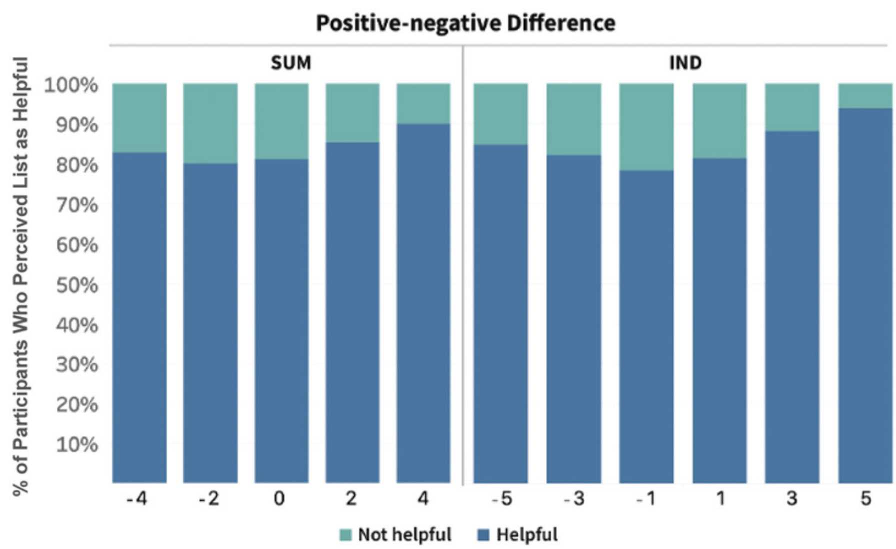

Figure 4. Effect of the valence of majority opinion on perceived helpfulness.

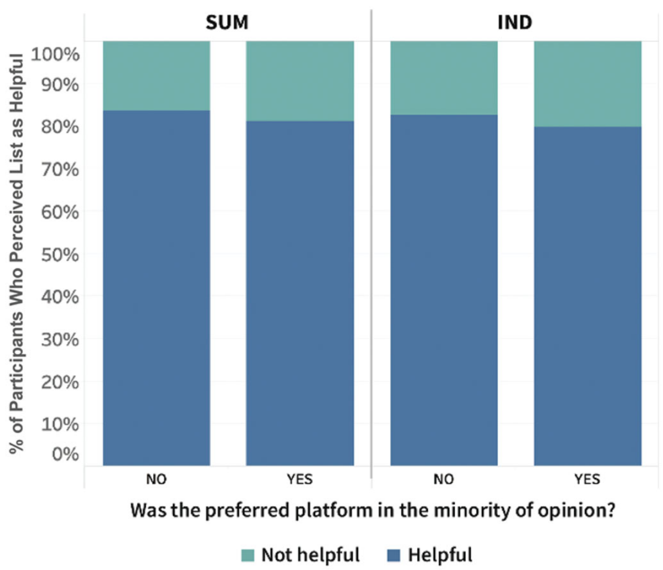

Figure 5. Main effects of preferred platforms' minority-opinion status on lists' perceived helpfulness: significant in IND, but non-significant in SUM.

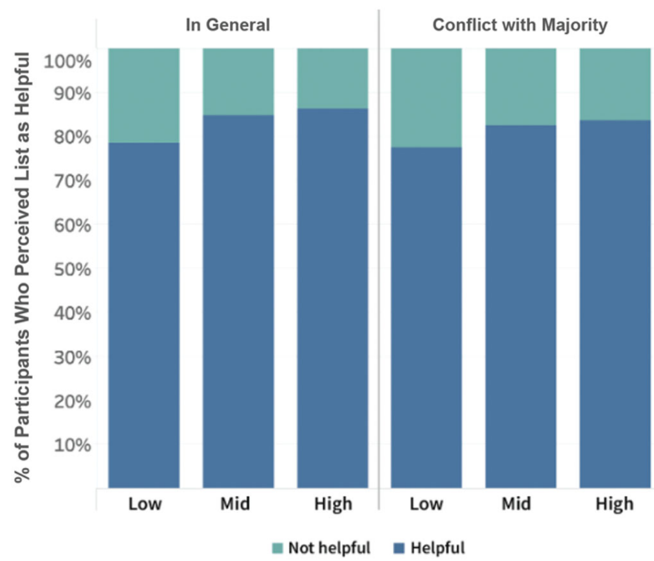

Figure 6. Effects of preferred-source quantity on lists' perceived helpfulness.

(RQ2-1), we found that, as expected, overall perceived list helpfulness was higher when such quantity was rated mid (85.18%) or high (87.25%) than when it was low (79.88%). Both these differences were statistically significant (low vs. mid: $Z = 5.353$, $p < .001$; low vs. high: $Z = 7.26$, $p < .001$). However, the difference between mid and high was only marginal ($Z = 1.706$, $p = .0879$). This is in accordance with prior findings that higher review quantity led users to perceive ratings as more credible, and thus more helpful (Flanagin & Metzger, 2013). Moreover, it indicates that the influence of participants' review-platform preferences was generally strong enough to affect the perceived helpfulness of the entire list.

We further looked into the effect of review quantity when the preferred platform expressed a minority opinion. As mentioned above, when such quantity is higher, users generally perceive aggregated ratings as more credible; thus, we expected that a high quantity of reviews on their preferred platform would lead our participants to perceive starker conflicts between that platform and the others, resulting in lower perceived helpfulness of the lists users were shown. Our participants indeed perceived SUM lists to be slightly less helpful when the quantity of reviews on their preferred platform was mid (78.79%) or high (81.89%) than when such quantity was low (82.63%), as can be seen from the right-hand side of Figure 6; and both these differences were statistically significant (low vs. mid: $Z = -3.976$, $p < .001$; low vs. high: $Z = -4.353$, $p < .001$). However, this lower perceived list helpfulness seems to have had a ceiling effect: that is, we did not observe a statistically significant difference in perceived helpfulness between mid and high quantity ($Z = -.344$, $p = .7307$). Because the number of reviews represented by the 'high' construct occupied a very wide range (from 96 to 3703), we also checked perceived helpfulness for quantity thresholds higher than 96. This indicated no increase in perceived helpfulness for any quantity above 400, for which it was 81.58%, in line with the

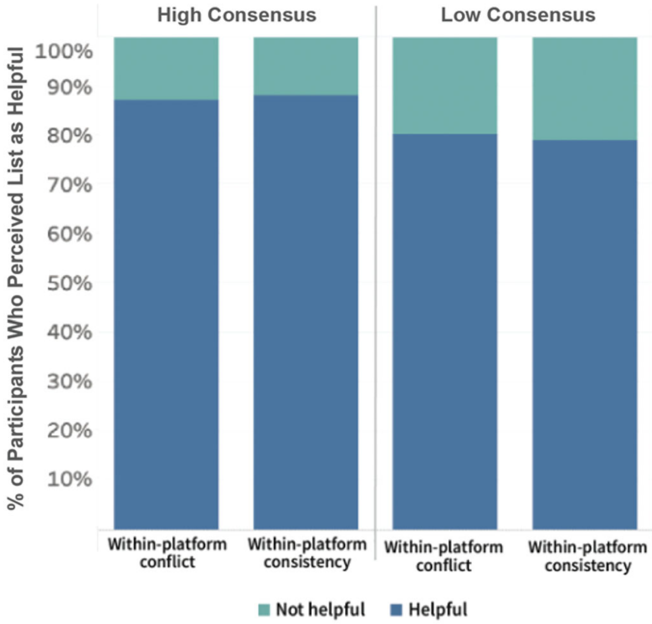

Figure 7. Effects of within-platform conflict vs. overall consensus on lists' perceived helpfulness.

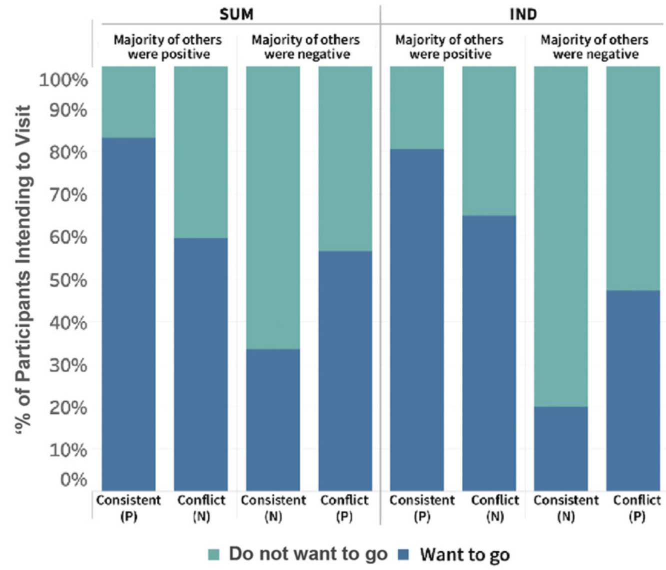

Figure 8. Differences in participants' visit intentions associated with opinion conflicts between their preferred platform and the majority of other platforms in a list.

all-quantities average. Thus, it seems that perceived helpfulness did not continue decreasing as review quantity increased from medium to high; so, as long as the quantity exceeded the threshold, this effect would probably be about the same.

Even more interesting, from our perspective, was that—so long as the quantity of reviews on their preferred platform was low—the participants actually deemed a given list to be significantly *more* helpful when that platform expressed a minority opinion (82.63%) than when it did not (79.88%). One explanation may be that these participants did not think a platform with a low quantity of reviews, despite being preferred, was persuasive enough to constitute a conflicting opinion, and thus did not feel it decreased list helpfulness. In other words, perceived list helpfulness in cases of opinion conflict appears to be dependent on review quantity.

4.1.4. Between-platform vs. Within-platform conflict

We next examined the influence of within-platform conflict in IND lists on perceived list helpfulness. For this purpose, we deemed consensus ratios of 3:2 to be low-consensus, and 4:1 and 5:0 to be high-consensus. As shown in Figure 7, perceived list helpfulness was significantly higher in a high-consensus condition than in a low-consensus one, regardless of whether a within-platform conflict was present (High-consensus: Within-platform conflict 87.03% vs. Platform consistent 87.95%; Low-consensus: Within-platform conflict 78.83% vs. Platform consistent: 80.09%). However, within-platform conflict and consistency were not associated with any significant differences in perceived helpfulness, irrespective of lists' overall consensus. This result appears to reflect the high strength of the main effect of between-platform conflict ($Z = -16.761$, $p < .001$), and the lack of any main effect of within-platform conflict, either in general ($Z = -.389$, $p = .697$) or with reference to preferred platforms in particular ($Z = -.237$, $p = .813$). In other words, the main drivers of reductions in lists' perceived helpfulness were whichever platforms had caused overall consensus to be low.

4.2. Influence of opinion conflicts on visit intention

4.2.1. Opinion conflicts between the preferred platform and others

Having ascertained the participants' perceptions of each list's helpfulness, we examined which opinion they accepted when their preferred platform's opinion conflicted with the majority of other platforms' opinions. To this end, we separately observed the visit-intention outcomes of review situations in which the majority of non-preferred sources were positive vs. negative, because these two valences tend to affect such intention in opposite directions. As shown in Figure 8, when the majority of other platforms expressed positive opinions, a negative opinion from the participant's preferred platform resulted in significantly lower visit intention: i.e., 59.65% (vs. 83.17%) in the case of SUM lists, and 64.93% (vs. 80.48%) in the case of IND ones. These differences were highly statistically significant (SUM: $Z = -16.69$, $p < .001$; IND: $Z = -16.95$, $p < .001$), and yet, visit intention remained above the 50% level in all cases. Likewise, when the majority of platforms expressed a negative opinion, the preferred platform's positive opinion led to considerably higher visit intention, i.e., 56.45% (vs. 33.49%) for SUM, and 47.11% (vs. 19.92%) for IND; and these differences were also both highly statistically significant (SUM: $Z = 14.177$, $p < .001$; IND: $Z = 16.59$, $p < .001$). From this, we can see that preferred-platform opinion profoundly influenced visit intention even when it conflicted with majority opinion. On the other hand, the same results show that participants typically agreed with platforms' majority opinion: i.e., their preferred

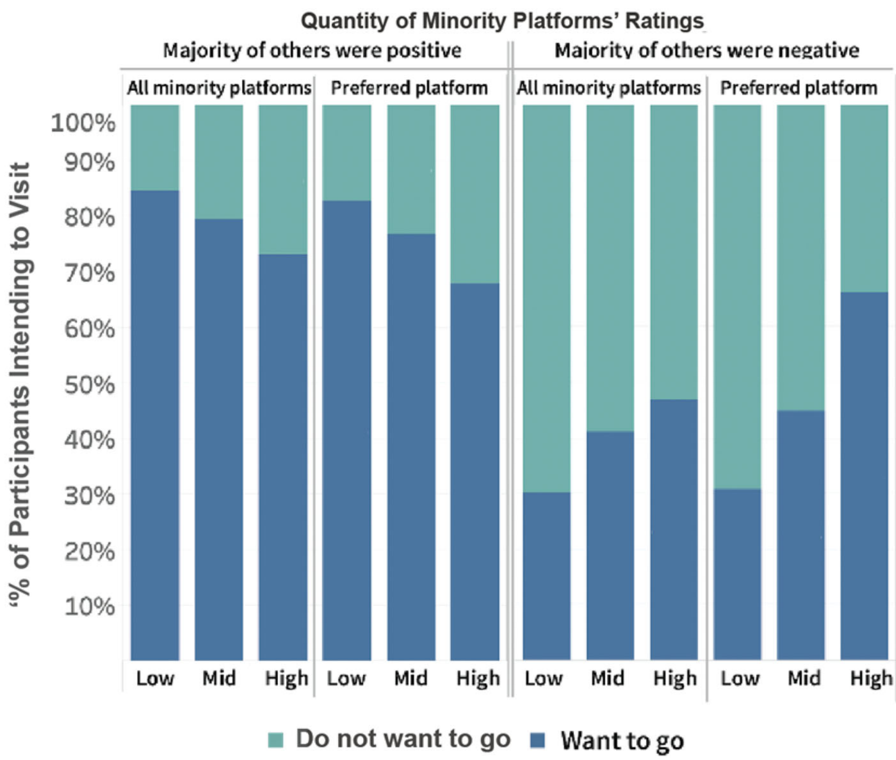

Figure 9. Effect of minority-opinion platforms on visit intention, by valence and review quantity.

platform did not change the valence of their visit intention in three out of the four studied conflict conditions. The one exception was when their preferred platform held the minority positive opinion in a SUM list: a condition associated with visit intention of more than 50%. The results reported in the next section may help explain this apparent anomaly.

4.2.2. Quantity of preferred-platform reviews

When the participant's preferred platform held the minority negative opinion within its list (Figure 9, far left), higher review quantity on that platform was associated with lower restaurant-visit intentions (i.e., low: 82.66%, mid: 76.69%, high: 67.85%). However, only the difference between high and low quantity, not between mid and low quantity, was statistically significant in this case ($Z = -3.127$, $p = .0017$). This finding suggests that the influence of the preferred platform on visit intention operated mainly above a certain quantity threshold. More than 50% of the time, participants remained willing to visit particular restaurants despite their preferred platforms' minority negative opinions of those restaurants, suggesting that they primarily followed majority opinion. However, when their preferred platform expressed a minority *positive* opinion (Figure 9, second from right), a higher quantity of preferred-source reviews was associated with significantly higher visit intention: more than 66% when such quantity was high, or more than double their visit intention when it was low (i.e., 30.91%). All differences between quantity levels were statistically significant (mid vs. low: $Z = 2.424$, $p = .0153$; high vs. low: $Z = 6.007$, $p < .001$; high vs. mid: $Z = 4.013$, $p < .001$). This indicates that the influence of the preferred platform's minority positive opinion was strongly dependent on its review quantity being above a certain threshold, and that when such quantity was high, the preferred platform's influence was strong enough to change the valence of restaurant-visit intentions. This also explains why, in a general case where the preferred platform held a minority positive opinion in a SUM list, visit intention remained above 50%.

We further examined whether the pattern observed above was specific to preferred platforms. As can be seen from Figure 9 (second from left), visit intention primarily followed majority opinion when any platform expressed a minority negative opinion, irrespective of preferred-platform status. However, as Figure 9 (far right) shows, the impact on visit intention of a positive rating by a non-preferred minority-opinion platform was markedly less than the impact of its preferred-platform counterpart.

4.2.3. Between-platform vs. Within-platform conflict

We only examined instances of the 3:2 consensus ratio (i.e., 3P2N and 2P3N), because restaurant-visit intention is largely affected by opinion valence; that is, in any review list with a 4:1 or 5:0 consensus ratio, such intention will be predominantly affected by majority positive or negative opinion, making it difficult to isolate the influence of within-platform conflict. Figure 10 (left) illustrates participants' visit intention with a consensus ratio of 3:2 in a general case (i.e., regardless of whether or not the platform that appeared twice was a preferred one). As expected, our participants expressed a much higher visit intention in 3P2N than in 2P3N. However, no significant differences in visit intention, in either 3P2N or 2P3N, were associated with the twice-appearing source being 2 P vs. 1P1N vs. 2 N. This suggests that participants' visit

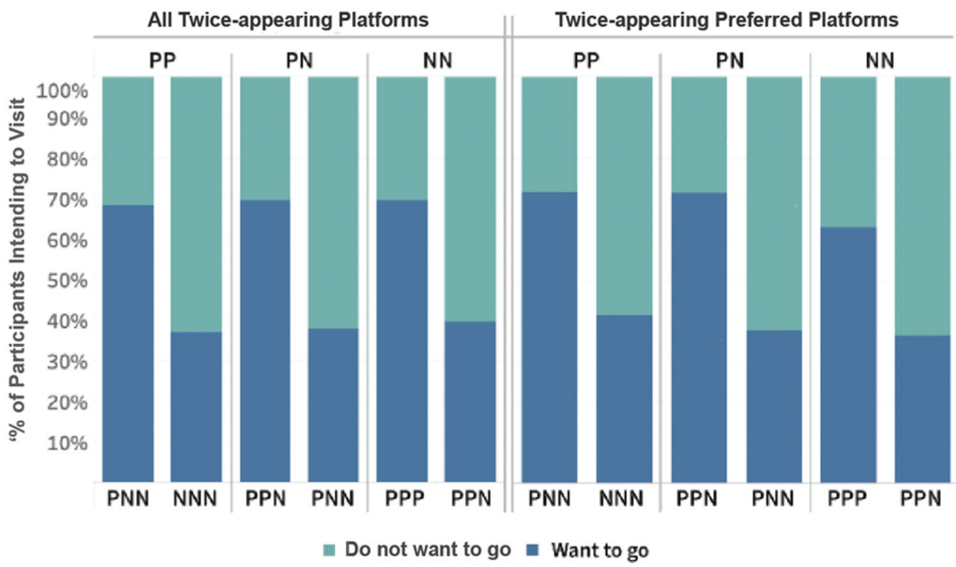

Figure 10. Effect on visit intention of all platforms to appear twice within a list (left) vs. preferred platforms that appeared twice (right).

intention was mainly affected by the proportion of positive vs. negative reviews in the list, as opposed to the presence of conflicting opinions from the same platform.

Next, we examined situations in which the twice-appearing platform was the participant's preferred one. However, as Figure 10 (right) shows, visit intention was again mainly affected by the overall proportion of positive vs. negative reviews. In short, within-platform conflicts did not affect visit intention.

5. Discussion

5.1. Opinion conflicts between the preferred platform and others

When the consensus restaurant opinion expressed by a participant's preferred platform conflicted with those of the majority of other platforms in the same list, there were two general outcomes: (1) the list was perceived as less helpful, and (2) the majority opinion was adopted. This result is encouraging, as it suggests that the strength of media bias in the sphere of experience-product reviews is not as strong as it can be in news reading (An et al., 2013; Stroud, 2011), where users are highly selective about the media they trust. These results highlight the considerable utility of exposing users to multi-platform lists of experience-product ratings and/or reviews, insofar as people are not persuaded solely by the platforms they already prefer when it comes to experience-product purchasing intentions. However, this is not to say that media bias does not exist in such scenarios. Indeed, the platforms on which our participants preferred to search for restaurant information still mattered to their visit intentions. For instance, when there was an obvious conflict of opinion about a given restaurant between a preferred platform and others, the preferred platform's positive (negative) opinion increased (decreased) the participants' overall willingness to visit it. In addition, the participants sometimes perceived a list as less helpful when such conflicts of opinion were present in it. Nevertheless, such differences in lists' perceived helpfulness were fairly small, because the participants took all platforms' opinions into account.

5.2. Moderating effect of preferred-platform review quantity

However, there were two main exceptions to these general outcomes, both of which were related to preferred-platform review quantity. The first was that a review list's perceived helpfulness was not damaged by the presence of conflicting opinions, provided that the participants' preferred platforms' review quantity was low. One prior study (Flanagin & Metzger, 2013) indicated that such quantity can act as a warrant, such that people generally perceive summarized or aggregated ratings based on high review quantities as more credible and reliable than those based on low ones. This could help explain why our participants did not find their preferred platforms' outlier opinions reliable: i.e., they did not see conflicts of opinion as genuine when their preferred platform's consensus opinion appeared to be endorsed by only a few people. As review quantity increased, however, the perceived persuasiveness of the preferred platform's opinion increased, making such conflicts more noteworthy. This, in turn, heightened perceptions of a genuine conflict between two opinions, and thus reduced the list's perceived helpfulness.

The other main exception was the particularly strong impact of a preferred platform's consensus restaurant opinion when the quantity of reviews it was based upon was high: to the point that two-thirds of the time, participants simply adopted the opinions of their preferred platforms. However, this strong influence was not observed when the preferred platform expressed a minority negative opinion. Prior research has suggested that individuals tend to be more susceptible to the influence of negative reviews than positive ones (Purnawirawan et al., 2014). However, our findings indicate that when using a multi-platform summarized rating list, people may be particularly susceptible to a positive consensus opinion based on a high quantity of reviews and emanating from a preferred source. In light of this dynamic, it is not particularly surprising that we did not observe such a strong influence of preferred platforms in IND lists, in which the preferred platform's minority positive opinion—which was not underwritten by quantity metrics, or endorsed by anyone other than its individual writer—never resulted in visit intentions above 50%. Together, these results indicate that, in SUM lists, the influence of preferred-platform minority opinion could be moderated by review quantity. That is, having a low quantity of reviews from the preferred platform did not lower such lists' perceived helpfulness, but if the quantity on which that platform's positive opinion was based was high, participants' restaurant-visit intention was significantly high.

Importantly, reviews in IND lists were not associated with quantity information, and thus, each could only represent one person's opinion, making their reliability and credibility questionable, especially when in conflict with reviews from the list's other platforms. Because quantity information served as an important basis both for individuals' perceptions of list helpfulness and their visit intentions, the choice of an IND vs. SUM list format sometimes led to differential user outcomes, despite the various platforms' consensus opinions being the same.

5.3. Between-platform vs. Within-platform conflict

Prior research on the influence of review consensus within a single review platform (e.g., Baek et al., 2012; Qiu et al., 2012; Quaschning et al., 2015), indicated that the perceived helpfulness of reviews of a product is lower if they conflict with majority opinion. In the case of review lists combining opinions from multiple platforms, we likewise found that the higher the consensus level within a list was, the more helpful people perceived that list to be. We expected that within-platform conflict would reduce such perceived helpfulness, but in fact found that perceptions of a multi-platform list as unhelpful were mainly driven by conflicts of opinion between its different platforms.

Conceivably, within-platform conflict still had some influence, which was masked reviews from multiple platforms were shown—perhaps because the participants mainly focused on consensus levels among disparate review platforms when making their restaurant-visit judgments. Therefore, in a multi-platform review list, the opinion of every platform counts.

5.4. Design implications

The findings of this study have some important design implications for future services offering multi-platform review lists. For one, they imply that each opinion in such lists counts, because users perceive each list as a whole and take account of its majority opinion when making purchase decisions. The slightly lower perceived helpfulness of such lists when they contain conflicts of opinion can probably be counterbalanced by the benefits of presenting reviews from diverse perspectives. Thus, we recommend that developers include reviews from multiple platforms, and that they not filter review lists simply to avoid the presentation of conflicting opinions. Indeed, the presence of such conflicts may tend to allay users' concerns that such lists have been subjected to commercial manipulation.

Based on our results, we believe that multi-platform review-list services could foster a more competitive atmosphere that encourages all platforms to provide more professional, higher-quality information. Because conflicts of opinion between different platforms are unavoidable, and because every source matters in a multi-sourced list, we recommend that such services favor sources with high influence, good reputation, and strong credibility, to (1) provide their users with reviews perceived as having strong authenticity and value, and (2) themselves become more resistant to malicious tampering.

Additionally, to balance the influence and fairness of review sources and encourage lesser-known, yet professional sources, such services should consider giving preferential weights to relatively small-scale sources, as this would enhance both the selectivity and diversity of the reviews they host. Such weightings could be adjusted rapidly according to users' credibility perceptions, recognition, and usage rates of each, to avoid a few well-known sources becoming dominant.

Lastly, each web user tends to have an established preference when it comes to sourcing review information, and his/her impression of an experience product may be profoundly influenced by a preferred platform's consensus opinion, especially when that consensus is based on a high quantity of discrete reviews or ratings. Thus, we recommend that future services hosting multi-platform review lists ascertain each user's preferred platform, through his/her past usage behaviors and/or direct questioning, and then identify whether s/he is mainly influenced by that platform or by majority opinion, as measured by actual restaurant-visit check-ins. And, given that our participants were likely to prefer different platforms depending on their assigned dining purposes, such list services should also include more comprehensive arrays of dining scenarios, and ask their users about scenario-dependent variation in their preferred review platforms.

5.5. Limitations and directions for future work

This study is subject to a number of limitations. First, its survey data is not internationally representative, since MTurk members only come from certain countries (Difallah et al., 2018). Additionally, existing restaurant-review platforms can be quite country-specific, e.g., Yelp in the U.S., TheFork in Europe, Dianping in China, etc.; but no such platforms were included in our study, due to the participants' likely unfamiliarity with some of them. Thus, our study's outcomes might have been different if it had included some of these specialized restaurant platforms, and future researchers should consider doing so. We also chose a single category of experience product, i.e., restaurants, so whether our results are applicable to other categories of such products (e.g., movies and hotels) will require further investigation.

In terms of our experimental stimuli, in both SUM and IND lists, we did not provide photos with reviews to avoid the influence of photo content, which was not the main focus of this study. In real life, however, people making decisions about restaurants are likely to browse photos of its dining environment and/or food.

Also, our IND lists only showed participants five reviews at a time, whereas in real life, people can view many more; and, because only one platform ever appeared twice in a given list, there was only one possible type of within-platform conflict, i.e., 1P1N. Thus, our results might have differed if 2P1N, 2N1P, or even more variants of within-platform conflict had been provided. And, because we attempted to control the influence of textual comments within each individual review, the comments used in the experiment were not fully representative of real-life situations. Given that machine-learning-based sentiment analysis (or "opinion mining") has long been applied to analysis of people's textual reviews (Feldman, 2013), it has been suggested that an efficient classification scheme could be a powerful tool for evaluating and predicting the performance of certain products or services, e.g., educational platforms, teaching, and restaurants (Gan et al., 2017; Onan, 2020, 2021; Zahoor et al., 2020). As such, we recommend that future researchers create a more ecologically valid platform and look into the impact of free-form comments on viewers of multi-platform review lists.

Looking beyond experimental stimuli, our study sought to minimize the ordering effect of sources by randomly determining ratings/reviews' order each time they were presented to the participants. However, this does not mean that such order had no impact at all. And last but not least, we analyzed perceived helpfulness and visit intention based on the participants' responses to the list as a whole, and did not consider their separate reactions to each list item. Future research should therefore do so.

6. Conclusion

This study is believed to be the first to investigate multi-platform review lists' influence on their users' perceptions of restaurants: a category of experience products about which information is very commonly sought via mobile local searches (Teevan et al., 2011). Our research has yielded some novel findings. First, when the opinion expressed by his/her preferred review platform conflicted with the majority opinion of the other platforms in a list, the user typically adopted the list's majority opinion, though the preferred platform's viewpoint still influenced his/her visit intentions. Interestingly, in such situations, the influence of the preferred platform differed across our two types of lists, i.e., INDs (individual review lists) and SUMs (summarized lists), because in the former, each review represented only one person's opinion, whereas the latter incorporated information on how many users had contributed to the average/consensus rating. Provided that such user contributions in SUMs were above a certain quantity threshold, the preferred platform was particularly influential on our participants' restaurant-visit intentions. On the other hand, when the quantity of user ratings that fed into the consensus opinion of a person's preferred platform was too low to be persuasive, conflicts of opinion between that platform and the others in a list were not influential enough to lower his/her perceptions of that list's helpfulness. In addition, low perceived list helpfulness could mainly be ascribed to the conflicts between platforms, while within-platform conflict in IND lists had little impact.

These results imply that presenting multi-platform lists of restaurant-review information is a promising approach for practitioners, insofar as web users do not simply adopt the opinions of the platforms they prefer. Although we found that within-list opinion conflicts did result in lower perceived list helpfulness, as compared to when such conflicts were absent or weak, the perceived helpfulness of our lists was generally high. Through this study, we have gained a preliminary understanding of how web users perceive multi-platform review lists when making restaurant-visit decisions. On that basis, we have offered a number of recommendations to developers of future multi-platform review-list services, in the hope that such services will take account of their users' individual preferences and behaviors, thus not only enhancing user experience, but also helping people select restaurants more efficiently.

Notes

1. https://en.wikipedia.org/wiki/LocalWiki.
2. https://maps.google.com/localguides/home

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

We sincerely thank the Ministry of Science and Technology, R.O.C. for their support [MOST 108-2218-E-009-050].

ORCID

Yung-Ju Chang http://orcid.org/0000-0001-6956-3459

References

An, J., Quercia, D., & Crowcroft, J. (2013). Fragmented social media: A look into selective exposure to political news. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 51–52). https://doi.org/10.1145/2487788.2487807

Bae, S., & Lee, T. (2011). Product type and consumers' perception of online consumer reviews. *Electronic Markets*, *21*(4), 255–266. https://doi.org/10.1007/s12525-011-0072-0

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, *17*(2), 99–126. https://doi.org/10.2753/JEC1086-4415170204

Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, *15*(3), 31–40. https://doi.org/10.1002/dir.1014

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, *37*(1), 48–74. https://doi.org/10.1037/0022-3514.37.1.48

Brown, J., Broderick, A. J., & Lee, N. (2007). Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of Interactive Marketing*, *21*(3), 2–20. https://doi.org/10.1002/dir.20082

Carbon, K. C., & Stanford, A. (2014). *Applications of machine learning to predict yelp ratings*. Stanford University.

Carlson, K., Guha, A., & Daniels, R. M. (2011). The ratings paradox: Why we prefer reading negative reviews, but then subsequently rate these reviews as less useful. *ACR North American Advances.*-NA, 38. https://www.acrwebsite.org/volumes/16359/volumes/v38/NA-38

Chang, C. (2011). Opinions from others like you: The role of perceived source similarity. *Media Psychology*, *14*(4), 415–441. https://doi.org/10.1080/15213269.2011.620539

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In *Dual-process theories in social psychology* (pp. 73–96). Guilford Press.

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, *54*(3), 477–491. https://doi.org/10.1287/mnsc.1070.0810

Cheng, Y.-H., & Ho, H.-Y. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research*, *68*(4), 883–887. https://doi.org/10.1016/j.jbusres.2014.11.046

Cheung, C., Sia, C. L., & Kuan, K. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems*, *13*(8). https://doi.org/10.17705/1jais.00305

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, *13*(4), 9–38. https://doi.org/10.2753/JEC1086-4415130402

Choi, J. W., & Ok, C. (2011). The effect of online restaurant reviews on diners' visit intention: a comparative analysis of expert vs. peer reviews. *In Graduate Student Research Conference in Hospitality and Tourism*. https://scholarworks.umass.edu/gradconf_hospitality/2011/Poster/39

Dabholkar, P. A. (2006). Factors influencing consumer choice of a "rating web site": An experimental investigation of an online interactive decision aid. *Journal of Marketing Theory and Practice*, *14*(4), 259–273. https://doi.org/10.2753/MTP1069-6679140401

De Pelsmacker, P., & Janssens, W. (2007). A model for fair trade buying behaviour: The role of perceived quantity and quality of information and of product-specific attitudes. *Journal of Business Ethics*, *75*(4), 361–380. https://doi.org/10.1007/s10551-006-9259-2

Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and dynamics of mechanical Turk workers. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 135–143). https://doi.org/10.1145/3159652.3159661

Doh, S.-J., & Hwang, J.-S. (2009). How consumers evaluate eWOM (electronic word-of-mouth) messages. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *12*(2), 193–197. https://doi.org/10.1089/cpb.2008.0109

Dou, X., Walden, J. A., Lee, S., & Lee, J. Y. (2012). Does source matter? Examining source effects in online product reviews. *Computers in Human Behavior*, *28*(5), 1555–1563. https://doi.org/10.1016/j.chb.2012.03.015

EdD, S. G., & MS, J. K. (2004). Restaurant choice. *Journal of Foodservice Business Research*, *7*(1), 81–95. https://doi.org/10.1300/J369v07n01_06

Fan, Y. C., & Chen, H. Y. (2014). Design and implementation of a real-time geographical query answering system. In *2014 International Conference on Information Science, Electronics and Electrical Engineering* (Vol. 2, pp. 1221–1225). https://doi.org/10.1109/InfoSEEE.2014.6947864

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89. https://doi.org/10.1145/2436256.2436274

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261–1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Filieri, R., Hofacker, C. F., & Alguezaui, S. (2018). What makes information in online consumer reviews diagnostic over time? The role of review relevancy, factuality, currency, source credibility and ranking score. *Computers in Human Behavior*, *80*, 122–131. https://doi.org/10.1016/j.chb.2017.10.039

Filieri, R., Lin, Z., Pino, G., Alguezaui, S., & Inversini, A. (2021). The role of visual cues in eWOM on consumers' behavioral intention and decisions. *Journal of Business Research*, *135*, 663–675. https://doi.org/10.1016/j.jbusres.2021.06.055

Filieri, R., & McLeay, F. (2014). E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of Travel Research*, *53*(1), 44–57. https://doi.org/10.1177/0047287513481274

Filieri, R., McLeay, F., Tsui, B., & Lin, Z. (2018). Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services. *Information & Management*, *55*(8), 956–970. https://doi.org/10.1016/j.im.2018.04.010

Filieri, R., Raguseo, E., & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, *77*, 333–341. https://doi.org/10.1016/j.ijhm.2018.07.013

Flanagin, A. J., & Metzger, M. J. (2013). Trusting expert- versus user-generated ratings online: The role of information volume, valence, and consumer characteristics. *Computers in Human Behavior*, *29*(4), 1626–1634. https://doi.org/10.1016/j.chb.2013.02.001

Gan, Q., Ferns, B. H., Yu, Y., & Jin, L. (2017). A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism*, *18*(4), 465–492. https://doi.org/10.1080/1528008X.2016.1250243

Goldsmith, R. E., & Horowitz, D. (2006). Measuring motivations for online opinion seeking. *Journal of Interactive Advertising*, *6*(2), 2–14. https://doi.org/10.1080/15252019.2006.10722114

Hao, Y., Ye, Q., Li, Y., & Cheng, Z. (2010). How does the valence of online consumer reviews matter in consumer decision making? Differences between search goods and experience goods. In *2010 43rd Hawaii international conference on system sciences* (pp. 1–10). https://doi.org/10.1109/HICSS.2010.455

Harkins, S. G., & Petty, R. E. (1987). Information utility and the multiple source effect. *Journal of Personality and Social Psychology*, *52*(2), 260–268. https://doi.org/10.1037/0022-3514.52.2.260

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. https://doi.org/10.1007/s10683-011-9273-9

Huang, J.-H., & Chen, Y.-F. (2006). Herding in online product choice. *Psychology and Marketing*, *23*(5), 413–428. https://doi.org/10.1002/mar.20119

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, *59*(1), 19–39. https://doi.org/10.1111/j.1460-2466.2008.01402.x

(Jimmy) 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. *International Journal of Hospitality Management*, *30*(1), 178–183. https://doi.org/10.1016/j.ijhm.2010.04.008

Kanouse, D. E., & Hanson, L. R. Jr., (1987). Negativity in evaluations. In *Attribution: Perceiving the causes of behavior* (pp. 47–62). Lawrence Erlbaum Associates, Inc.

Khare, A., Labrecque, L. I., & Asare, A. K. (2011). The assimilative and contrastive effects of word-of-mouth volume: An experimental examination of online consumer ratings. *Journal of Retailing*, *87*(1), 111–126. https://doi.org/10.1016/j.jretai.2011.01.005

Lee, S.-H. (2009). How do online reviews affect purchasing intention? *African Journal of Business Management*, *3*(10), 576–581. https://doi.org/10.5897/AJBM09.204

Lee, C.-H., & Cranage, D. A. (2014). Toward understanding consumer processing of negative online word-of-mouth communication: The roles of opinion consensus and organizational response strategies *Journal of Hospitality & Tourism Research*, *38*(3), 330–360. https://doi.org/10.1177/1096348012451455

Li, M., Huang, L., Tan, C.-H., & Wei, K.-K. (2013). Helpfulness of online product reviews as seen by consumers: Source and content features. *International Journal of Electronic Commerce*, *17*(4), 101–136. https://doi.org/10.2753/JEC1086-4415170404

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89. https://doi.org/10.1509/jmkg.70.3.074

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*, 140–151. https://doi.org/10.1016/j.tourman.2014.09.020

Lu, L., Wu, L., & He, Z. (2020). Is your restaurant worth the risk? A motivational perspective on reviews' rating distribution and volume. *Journal of Hospitality & Tourism Research*, *44*(8), 1291–1317. https://doi.org/10.1177/1096348020944537

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200. https://doi.org/10.2307/20721420

Noone, B. M., & Mattila, A. S. (2010). Consumer goals and the service encounter: Evaluating goal importance and the moderating effect of goal progress on satisfaction formation. *Journal of Hospitality & Tourism Research*, *34*(2), 247–268. https://doi.org/10.1177/1096348009350619

Onan, A. (2020). Mining opinions from instructor evaluation reviews: A deep learning approach. *Computer Applications in Engineering Education*, *28*(1), 117–138. https://doi.org/10.1002/cae.22179

Onan, A. (2021). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. *Concurrency and Computation: Practice and Experience*, *33*(23), e5909. https://doi.org/10.1002/cpe.5909

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing*, *87*(4), 598–612. https://doi.org/10.1016/j.jretai.2011.05.002

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419.

Park, D.-H., & Kim, S. (2008). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, *7*(4), 399–410. https://doi.org/10.1016/j.elerap.2007.12.001

Park, D.-H., & Lee, J. (2008). EWOM overload and its effect on consumer behavioral intention depending on consumer involvement. *Electronic Commerce Research and Applications*, *7*(4), 386–398. https://doi.org/10.1016/j.elerap.2007.11.004

Park, D.-H., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125–148. https://doi.org/10.2753/JEC1086-4415110405

Park, S., Kang, S., Chung, S., & Song, J. (2009). NewsCube: Delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 443–452). https://doi.org/10.1145/1518701.1518772

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, *50*, 67–83. https://doi.org/10.1016/j.annals.2014.10.007

Pentina, I., Bailey, A. A., & Zhang, L. (2018). Exploring effects of source similarity, message valence, and receiver regulatory focus on yelp review persuasiveness and purchase intentions. *Journal of Marketing Communications*, *24*(2), 125–145. https://doi.org/10.1080/13527266.2015.1005115

Ponnam, A., & Balaji, M. S. (2014). Matching visitation-motives and restaurant attributes in casual dining restaurants. *International Journal of Hospitality Management*, *37*, 47–57. https://doi.org/10.1016/j.ijhm.2013.10.004

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, *34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Purnawirawan, N., De Pelsmacker, P., & Dens, N. (2012). Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *Journal of Interactive Marketing*, *26*(4), 244–255. https://doi.org/10.1016/j.intmar.2012.04.002

Purnawirawan, N., Dens, N., & De Pelsmacker, P. (2014). Expert reviewers beware! The effects of review set balance, review source and review content on consumer responses to online reviews. *Journal of Electronic Commerce Research.-Long Beach, Calif*, *15*(3), 162–178.

Qiu, L., Pang, J., & Lim, K. H. (2012). Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence. *Decision Support Systems*, *54*(1), 631–643. https://doi.org/10.1016/j.dss.2012.08.020

Quaschning, S., Pandelaere, M., & Vermeir, I. (2015). When consistency matters: The effect of valence consistency on review helpfulness. *Journal of Computer-Mediated Communication*, *20*(2), 136–152. https://doi.org/10.1111/jcc4.12106

Racherla, P., & Friske, W. (2012). Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories. *Electronic Commerce Research and Applications*, *11*(6), 548–559. https://doi.org/10.1016/j.elerap.2012.06.003

Ribeiro, F. N., Henrique, L., Benevenuto, F., Chakraborty, A., Kulshrestha, J., Babaei, M., Gummadi, K. P. (2018, June 15). Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Twelfth International AAAI Conference on Web and Social Media*. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17878

Rieh, S. Y. (2015). Credibility and cognitive authority of information. In *Encyclopedia of Library and Information Sciences* (3rd ed.). CRC Press.

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, *41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Roy Dholakia, R., & Sternthal, B. (1977). Highly credible sources: Persuasive facilitators or persuasive liabilities? *Journal of Consumer Research*, *3*(4), 223–232. https://doi.org/10.1086/208671

Ruiz-Mafe, C., Chatzipanagiotou, K., & Curras-Perez, R. (2018). The role of emotions and conflicting online reviews on consumers' purchase intentions. *Journal of Business Research*, *89*, 336–344. https://doi.org/10.1016/j.jbusres.2018.01.027

Schindler, R., & Bickart, B. (2005). Published word of mouth: Referable, consumer-generated information on the internet. In C. P. Haugtvedt, K. A. Machleit, & R. F. Yalch (Eds.), *Online consumer psychology; understanding and influencing consumer behavior in the virtual world* (pp. 35–62). Lawrence Erlbaum Associates.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159–169. https://doi.org/10.1016/j.jretai.2004.04.001

Slater, M. D., & Rouner, D. (1996). How message evaluation and source attributes may influence credibility assessment and belief change. *Journalism & Mass Communication Quarterly*, *73*(4), 974–991. https://doi.org/10.1177/107769909607300415

Smith, D., Menon, S., & Sivakumar, K. (2005). Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing*, *19*(3), 15–37. https://doi.org/10.1002/dir.20041

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, *32*(6), 1310–1323. https://doi.org/10.1016/j.tourman.2010.12.011

Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, *39*, 1–9. https://doi.org/10.1016/j.tourman.2013.03.007

Stroud, N. J. (2011). *Niche news: The politics of news choice*. Oxford University Press.

Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, *14*(1), 47–65. https://doi.org/10.1287/isre.14.1.47.14767

Teevan, J., Karlson, A., Amini, S., Brush, A. J. B., & Krumm, J. (2011). Understanding the importance of location, time, and people in mobile local search behavior. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (pp. 77–80). https://doi.org/10.1145/2037373.2037386

Todorov, A., Chaiken, S., & Henderson, M. D. (2002). The Heuristic-systematic model of social information processing. In *The Persuasion handbook: Developments in theory and practice* (pp. 195–212). SAGE Publications. https://doi.org/10.4135/9781412976046

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134–144. https://doi.org/10.1002/asi.10016

Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, *5*(1), 15–28. https://doi.org/10.7903/cmr.1145

Willemsen, L. M., Neijens, P. C., Bronner, F., & de Ridder, J. A. (2011). "Highly recommended!" The content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication*, *17*(1), 19–38. https://doi.org/10.1111/j.1083-6101.2011.01551.x

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science*, *21*(2), 101–112. https://doi.org/10.1007/BF02894421

Wu, Y., Liu, T., Teng, L., Zhang, H., & Xie, C. (2021). The impact of online review variance of new products on consumer adoption intentions. *Journal of Business Research*, *136*, 209–218. https://doi.org/10.1016/j.jbusres.2021.07.014

Xue, F., & Phelps, J. (2004). Internet-facilitated consumer-to-consumer communication: The moderating role of receiver characteristics. *International Journal of Internet Marketing and Advertising*, *1*(2), 121. https://doi.org/10.1504/IJIMA.2004.004016

Yang, J., & Mai, E. (Shirley). (2010). Experiential goods with network externalities effects: An empirical study of online rating system. *Journal of Business Research*, *63*(9–10), 1050–1057. https://doi.org/10.1016/j.jbusres.2009.04.029

Yang, J., Sarathy, R., & Walsh, S. M. (2016). Do review valence and review volume impact consumers' purchase decisions as assumed? *Nankai Business Review International*, *7*(2), 231–257. https://doi.org/10.1108/NBRI-11-2015-0028

Zahoor, K., Bawany, N. Z., & Hamid, S. (2020). Sentiment analysis and classification of restaurant reviews using machine learning. In *2020 21st International Arab Conference on Information Technology (ACIT)* (pp. 1–6). https://doi.org/10.1109/ACIT50332.2020.9300098

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality*, *11*(3), 1–37. https://doi.org/10.1145/3309699

Zhang, C., & Conrad, F. (2014). Speeding in Web Surveys: The tendency to answer very fast and its association with straightlining. *Survey Research Methods*, *8*(2), 127–135. https://doi.org/10.18148/SRM/2014.V8I2.5453

Zhang, K., Lee, M., & Zhao, S. (2010). *Understanding the informational social influence of online review platforms* (Vol. 71). https://aisel.aisnet.org/icis2010_submissions/71

Zhang, K. Z. K., Zhao, S. J., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, *67*, 78–89. https://doi.org/10.1016/j.dss.2014.08.005

Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, *29*(4), 694–700. https://doi.org/10.1016/j.ijhm.2010.02.002

About the Authors

Tzu-Hao Lin received the M.S. degree in Computer Science, National Chiao Tung University (NCTU). He is currently working in Mediatek as a software engineer, mainly responsible for the development of automation tools and data analysis & visualization. His research interests mainly focused on human computer interaction on mobile devices.

Yen-Yun Liu received her B.S. degree in Psychology from National Taiwan University (NTU), Taipei, Taiwan. She is currently pursuing an M.Sc. degree in Applied Cognitive Psychology at Utrecht University, Utrecht, the Netherlands. Her research interests include mobile crowdsourcing, social robotics, and autonomous driving.

Hong-Han Shuai received the B.S., M.S. and Ph.D. degree from the Department of Electrical Engineering, National Taiwan University (NTU), Taiwan He is now an associate professor in National Yang Ming Chiao Tung University (NYCU), Taiwan. His research interests are multimedia processing, machine learning, social network analysis, and data mining.

Fang Hsin Hsu received the B.S. degree in Computer Science & Information Engineering, National Central University (NCU). She is currently studying for her M.S. degree at the Institute of Network Engineering at NYCU. Her research interest is human computer interaction (HCI), with a focus on attention research.

Yung-Ju (Stanley) Chang is an associate professor at the Department of Computer Science at NYCU. He received his M.S. and Ph.D. degree in Information Science from the University of Michigan. His research interest is human computer interaction (HCI), with a focus on attention research in computer-mediated communication.