CRUCIAL COMPONENTS: EXPLORING HOW REVIEW FORMATS AFFECT PURCHASES

Dr. Xiao Jun Wang[1], Prof. Yuan Liu2, Dr. Jing Hua Zhou[3]

Abstract

The escalating trend of online shopping has propelled the significance of online reviews as a potent influencer of consumer behavior. With a focus on the Chinese market, where online shopping has surged dramatically, this study examines the impact of online reviews on consumers' purchasing decisions. The proliferation of online shoppers, exemplified by 749 million individuals in China as of June 2020, underscores the pervasive influence of digital commerce. Notably, the pivotal role of online reviews is evident, with 77.5% of consumers engaging with these reviews, as reported by the China Internet Network Information Center (CNNIC). Moreover, empirical evidence from a survey of 140 consumers underscores that 91.43% of purchase determinations are contingent on both the volume and sentiment of reviews.

A distinctive observation pertains to the diverse presentation of online reviews on various e-commerce platforms. This study sheds light on this phenomenon, illustrating that platforms like Taobao.com and Tmall.com employ a frequency format (e.g., positive reviews: 1,672; total reviews: 1,823), while JD.com and Vip.com utilize a percentage format (e.g., positive review ratio: 91.7%; total reviews: 1,823). Through this investigation, the study unveils the intricate interplay between review presentation formats and consumers' decision-making processes in the context of online shopping.

---

[1] School of Management, Zhejiang University of Technology, Hangzhou, Zhejiang, China.

1   Introduction

The rate of online shopping has increased rapidly. In China, for example, in June 2020, the number of online shoppers had reached 749 million (China Internet Network Information Center [CNNIC], 2020). On the first day of the 2020 Double Eleven shopping promotion, online shopping transactions amounted to 498.2 billion yuan (about 77 billion US dollars; hangzhou.gov.cn, 2020). Online reviews are a key factor in influencing consumers' online shopping (De Pelsmacker et al., 2018; Nguyen et al., 2018). CNNIC (2016) reports that 77.5% of online consumers browse online reviews on the Internet. According to our survey of 140 randomly selected consumers, 91.43% decide whether to buy a product depending on review quantity and valence.

An important phenomenon that we observe is that daily online shopping platforms present online reviews in different formats. Some (Taobao.com and Tmall.com) use the frequency format (e.g., positive reviews: 1,672; total reviews: 1,823), while others (JD.com and Vip.com) use the percentage format (e.g., positive review ratio: 91.7%; total reviews: 1,823; see Figure 1).

Some scholars have found that when consumers make shopping decisions, they exhibit the *love of large numbers* (Powell et al., 2017). That is, consumers exhibit a strong bias favoring more-reviewed (and, thus, apparently more popular) products. On the one hand, the popularity of the product represents its quality to some extent (Chen, 2008). On the other hand, in accordance with the law of large numbers, if a product's evaluation is based on a large number of reviews, then it will be considered more reliable. But in the reality of online shopping, the presentation of online review information is very complex and diverse, involving frequency vs. percentage, positive and negative, many vs. few reviews, and so on. Do consumers always express a reliable preference for more-reviewed products? If consumers' purchase intentions are affected by the statistical format, then if platforms present online reviews in only one format, this may result in more perceptual bias and affect consumers' shopping choices. This may be unscientific.

Therefore, we conducted this research to test whether the recently proposed "*love of large numbers*" theory always exists and whether consumers have biases in the processing of online review information. Our goal is to provide guidelines for developing a more scientific and objective way of presenting online reviews.

In our research, we combined the presentations of review information on online shopping platforms in real situations. We found that the statistical information in online reviews involves four main variables: the statistical format, the review quantity, the review valence, and the review frame. Each variable, alone, may affect consumers' decisions, or the interaction of several variables at once may influence consumers' purchase intention. Therefore, the effect of each variable and the possible interaction effects of the four will be discussed in the following sections.

Figure 1

2   Theoretical Background and Hypotheses Development

2.1   Statistical Format: Frequency vs. Percentage

Frequency and percentage are the two forms of numerical representation that affect individuals' information focus, processing difficulty, and numerical evaluation (Dackermann et al., 2018; Romero et al., 2018). The numerator is more focused than the denominator in the frequency format, sometimes leading to a ratio bias in judgment (fuzzy-trace theory, Reyna, 2004). For example, patients *perceived* the cancer mortality rate (1286/10000) as higher than the actual cancer mortality rate (24.14/100) (Reyna & Brainerd, 2008). Investors may think nonproportionally — — Investors pay more attention to how much a stock has fallen, not the proportion (Shue, Townsend, 2021), in other words, they value absolute values (numbers) more than rates. Therefore, if the numerical information in product reviews is presented in the frequency format, because the total number of reviews is different for each product, the format will not help consumers make the most-informed judgments and choices. According to the basic assumption of the frequency hypothesis, humans have directly experienced frequencies, or counts, throughout their evolutionary history, making frequencies easier to understand compared to decimals or probabilities expressed on a 0 to 1.0 scale (i.e., normalized), which do not occur in nature (Brase, 2002). In general, information in the frequency format is more intuitive but harder to calculate and compare (Akl et al., 2011), whereas information in the percentage format is more accurate but more abstract (Waters et al., 2006).

2.2   Review Quantity

Consumers often evaluate products by relying on review quantity (Sotiriadis & Van Zyl, 2013): They perceive a larger review quantity as representing a more popular product (Powell et al., 2017) and as being associated with a higher demand (Zhu& Zhang, 2010) because a larger review quantity often represents social approval (Zhang et al., 2013).

2.3   Review Valence

Review valence refers to the proportion of negative (or positive) reviews to total reviews (Yang et al., 2016). It is a key indicator for perceiving the quality of products (Katz & Lazarsfeld, 2006). Many studies have reported that the higher the proportion of negative reviews, the worse consumers perceive the product to be and the lower the intentions of consumers to purchase the product (e.g., Lu et al., 2013).

2.4   Review Frame

Framing refers to the positive or negative description of an objective event (Kahneman & Tversky, 1979). Online reviews can be presented as a 90.6% applause ratio (positive frame) or as a 9.4% bad review ratio (negative frame).

Prospect theory shows an obvious framing effect: negative frames loom larger than positive frames (Kahneman & Tversky, 1979). In online shopping, researchers have observed that negative reviews have a greater impact on consumers than do positive reviews (Yang& Mai, 2010; Yin et al., 2014).

2.5   Statistical Format with Small vs. Large Review Quantity

In the frequency format, calculating the relative ratio (positive reviews/total reviews) occupies more cognitive load, leaving people with less energy or attention to process the total reviews. In the percentage format, the ratio is ready-made, and people have adequate energy to process the total reviews (Lee et al., 2019). The advantages and disadvantages of the total reviews are highlighted more in the percentage format than in the frequency format (Petrova et al., 2018). A large total of reviews should be perceived better in the percentage (vs. the frequency) format, and a small total of reviews should be perceived as worse in the percentage (vs. frequency) format. Therefore, we assume the following:

H1: The percentage format will induce higher purchase intentions than the frequency format with a large review quantity, whereas the frequency format will induce higher purchase intentions than the percentage format with a small review quantity.

2.6 Statistical Format with Positive vs. Negative Review Valence

The difference in the difficulty of processing frequency versus percentage may influence individuals' evaluation of information valence (McKechnie et al., 2012). The frequency format is more difficult to process than the percentage format, resulting in more distortion in valence evaluation (Lee et al., 2019). This makes the positive outcome less positive and the negative outcome less negative in the frequency (vs. percentage) format (Petrova et al., 2018). For instance, the advantage of positive review valence will be less obvious in the frequency format because of the difficulty of calculating the exact ratio. Similarly, the disadvantage of negative review valence will be less obvious in the frequency (vs. percentage) format due to the calculation difficulty. We hypothesize the following:

H2: The percentage format will induce higher purchase intentions than the frequency format in the positive review valence, whereas the frequency format will induce higher purchase intentions than the percentage format in the negative review valence.

2.7 Statistical Format with Positive vs. Negative Review Frame

According to fuzzy-trace theory, when reasoning and making decisions, people tend to rely on their memories to extract the essence of information, even when they can remember, verbatim, the details (e.g., quantitative) of such information (Reyna & Brainerd, 2008). So, in the frequency format, due to confusion created by overlapping or nested classes, people focus on the salient gist—often comparisons between numerators—at the expense of focusing on denominators (Reyna, 2004; Srivastava & Koukova, 2018). If the numerator information is more prominent, we speculate that in a positive frame, the frequency format (e.g., positive reviews: 22,400; total reviews: 23,300) would highlight the positive numerator information more than the percentage format (e.g., positive review ratio: 96.41%; total reviews: 23,300) would; however, in a negative frame, the frequency format (e.g., negative reviews: 900; total reviews: 23,300) would highlight the negative numerator information more than the percentage format (e.g., negative review ratio: 3.86%; total reviews: 23,300) would. The latter is more likely to arouse consumers' loss aversion, so that they overestimate the percentage of the number of negative reviews, thus leading to negative bias or valence distortion (Yang & Mai, 2010; Yin et al., 2014). Conversely, the percentage format allows consumers to get a more accurate estimate of the percentage of negative reviews. We hypothesize that the framing effect would be more obvious in the frequency format than in the percentage format.

H3: The frequency format will induce higher purchase intentions than the percentage format in a positive review frame, whereas the percentage format will induce higher purchase intentions than the frequency format in a negative review frame.

3 An Overview of the Current Research

We conducted two studies to test our hypotheses. In Study 1, we designed standard experimental scenarios, adopted a within-subjects design, and selected daily necessities, electronic products, and travel goods as materials. In Study 2, we designed emulation online shopping scenarios, adopted a between-subjects design, and selected clothing, food, and household appliances as materials to further test these hypotheses.

To detect a medium effect size of 0.25 at 95% power (α =.05), we ensured that there were at least 141 participants in Study 1 (within-subjects design), and 960 participants in Study 2 (between-subjects design).

4   Study 1

4.1   Participants, Materials, and Procedure

A 2 (statistical format: frequency vs. percentage) × 2 (review quantity: large vs. small) × 2 (review valence: positive vs. negative) × 2 (review frame: positive vs. negative) within-subjects experimental design was adopted to test our hypotheses. We selected three daily online products—daily necessities (shampoo), electronic products (a headset), and travel goods (a suitcase)—and provided online reviews, with 16 versions.

Consider the suitcase as an example.

Frequency, large quantity, positive valence, positive frame condition: Suitcase A: number of positive reviews: 14,922; total reviews: 15,423;

Percentage, large quantity, positive valence, positive frame condition: Suitcase B: percentage of positive reviews: 96.75%; total reviews: 15,423;

Frequency, large quantity, positive valence, negative frame condition: Suitcase C: number of negative reviews: 501; total reviews: 15,423;

Percentage, large quantity, positive valence, negative frame condition: Suitcase D: percentage of positive reviews: 3.25%; total reviews: 15,423;

Frequency, small quantity, positive valence, positive frame condition: Suitcase E: number of positive reviews: 149; total reviews: 154;

Percentage, small quantity, positive valence, positive frame condition: Suitcase F: percentage of positive reviews: 3.25%; total reviews: 154 (see Table 1 for other versions and product materials).

Table 1 *Experimental Materials (in Study 1)*

			Shampoo		Headset		Suitcase	
			Frequency format	Percentage format	Frequency format	Percentage format	Frequency format	Percentage format
Negative review valence	Large review quantity	Positive frame	The number of positive reviews: 76,111; the number of reviews: 85,124	The percentage of positive reviews: 89.41%; the number of reviews: 85,124	The number of positive reviews: 20,600; the number of reviews: 23,300	The percentage of positive reviews: 88.41%; the number of reviews: 23,300	The number of positive reviews: 13,110; the number of reviews: 15,423	The percentage of positive reviews: 85.00%; the number of reviews: 15,423
		Negative frame	The number of negative reviews: 9,013; the number of reviews: 85,124	The percentage of negative reviews: 10.59%; the number of reviews: 85,124	The number of negative reviews: 2,700; the number of reviews: 23,300	The percentage of negative reviews: 11.59%; the number of reviews: 23,300	The number of negative reviews: 2,313; the number of reviews: 15,423	The percentage of negative reviews: 15.00%; the number of reviews: 15,423
	Small review quantity	Positive frame	The number of positive reviews: 76; the number of reviews: 85	The percentage of positive reviews: 89.41%; the number of reviews: 85	The number of positive reviews: 206; the number of reviews: 233	The percentage of positive reviews: 88.41%; the number of reviews: 233	The number of positive reviews: 131; the number of reviews: 154	The percentage of positive reviews: 85.00%; the number of reviews: 154
		Negative frame	The number of negative reviews: 9; the number of reviews: 85	The percentage of negative reviews: 10.59%; the number of reviews: 85	The number of negative reviews: 27; the number of reviews: 233	The percentage of negative reviews: 11.59%; the number of reviews: 233	The number of negative reviews: 23; the number of reviews: 154	The percentage of negative reviews: 15.00%; the number of reviews: 154
Positive review valence	Large review quantity	Positive frame	The number of positive reviews: 83,120; the number of reviews: 85,124	The percentage of positive reviews: 97.65%; the number of reviews: 85,124	The number of positive reviews: 22,400; the number of reviews: 23,300	The percentage of positive reviews: 96.14%; the number of reviews: 23,300	The number of positive reviews: 14,922; the number of reviews: 15,423	The percentage of positive reviews: 96.75%; the number of reviews: 15,423
		Negative frame	The number of negative reviews: 2,004; the number of reviews: 85,124	The percentage of negative reviews: 2.35%; the number of reviews: 85,124	The number of negative reviews: 900; the number of reviews: 23,300	The percentage of negative reviews: 3.86%; the number of reviews: 23,300	The number of negative reviews: 501; the number of reviews: 15,423	The percentage of negative reviews: 3.25%; the number of reviews: 15,423
	Small review quantity	Positive frame	The number of positive reviews: 83; the number of reviews: 85	The percentage of positive reviews: 97.65%; the number of reviews: 85	The number of positive reviews: 224; the number of reviews: 233	The percentage of positive reviews: 96.14%; the number of reviews: 233	The number of positive reviews: 149; the number of reviews: 154	The percentage of positive reviews: 96.75%; the number of reviews: 154
		Negative frame	The number of negative reviews: 2; the number of reviews: 85	The percentage of negative reviews: 2.35%; the number of reviews: 85	The number of negative reviews: 9; the number of reviews: 233	The percentage of negative reviews: 3.86%; the number of reviews: 233	The number of negative reviews: 5; the number of reviews: 154	The percentage of negative reviews: 3.25%; the number of reviews: 154

The participants were asked to browse a series of products (randomly presented) with online review information and then to indicate their purchase intentions for these products (on a scale ranging from 1 = *very unwilling to buy* to 7 = *very willing to buy*). Cronbach's α =.93. Based on the calculated sample size given above, we selected 150 participants (64 females; M_{age} = 24.58, *SD* = 3.54) from Sojump (http://www.Wjx.cn), an online platform similar to Mechanical Turk or Qualtrics, which is used to launch nationwide e-surveys in China. We paid each participant ¥5 (¥1 = $0.14). Nine participants who did not pass the game rule comprehension test were excluded. To determine the thresholds for a small/large review quantity and a positive/negative review, we conducted a preliminary study (*N* = 100) to ask the participants (from the same formal study pool) to indicate the thresholds for a small and large review quantity and a positive and negative review for the shampoo, headset, and suitcase, respectively, based on their online shopping experience. We then calculated the mean and standard deviation of the reported thresholds and determined the formal experimental materials. Furthermore, we conducted a postcheck to test whether our thresholds matched the participants' experiences. For instance, the participants were asked the following questions: 1) Based on your online shopping experience for a shampoo, a total of 85,124 reviews is: A: large; B: small. 2) For a shampoo, 97.65% positive reviews and 2.35% negative reviews are: A: a positive valence; B: a negative valence (see Table 2 for analysis and Appendix A for all the post-check items).

Table 2 *The Post-Check Analysis (in Studies 1 and 2)*

		Percentage of positive reviews		Percentage of negative reviews		Number of reviews	
		High	Low	Low	High	Large	Small
Shampoo	Value setting	97.65%	89.41%	2.35%	10.59%	85124	85
	The percentage of people who agree	96.5%	70.9%	92.9%	79.4%	97.2%	95.7%
Headset	Value setting	96.14%	88.41%	3.86%	11.59%	23300	233
	The percentage of people who agree	94.3%	68.8%	92.2%	73.8%	94.3%	97.9%
Suitcase	Value setting	96.75%	85%	3.25%	15%	15423	154
	The percentage of people who agree	93.6%	74.5%	91.5%	76.6%	94.3%	97.2%
Hat	Value setting	98.28%	82.76%	1.72%	17.24%	11637	116
	The percentage of people who agree	99.3%	77.3%	97.9%	70.9%	94.3%	95.7%
Nuts	Value setting	99.07%	83.64%	0.93%	16.36%	42829	428
	The percentage of people who agree	99.3%	68.8%	97.9%	68.8%	97.2%	93.6%
TV	Value setting	97.24%	78.73%	2.76%	21.27%	36275	362
	The percentage of people who agree	96.5%	78.7%	93.6%	83.7%	95%	92.9%

4.2   Results

A 2 (statistical format: frequency vs. percentage) × 2 (review quantity: large vs. small) × 2 (review valence: positive vs. negative) × 2 (review frame: positive vs. negative) analysis of variance (ANOVA) was conducted. The results showed that the main effect of the review quantity was significant, $F(1, 140) = 19.40$, $p <.001$, η_p^2 =.12: the more reviews a product had, the more likely consumers were to buy it (M_{large}= 4.10, $SD = 1.62$, M_{small}= 3.74, $SD = 1.46$). The main effect of review valence was significant, $F(1, 140) = 473.01$, $p <.001$, η_p^2 =.77: when the review valence was positive, consumers were more likely to buy the product ($M_{positive}$= 4.59, $SD = 1.39$, $M_{negative}$= 3.25, $SD = 1.42$). The main effect of review frame was significant, $F(1, 140) = 303.93$, $p <.001$, η_p^2 =.69: participants' purchase intentions were higher in a positive frame ($M = 4.51$, $SD = 1.41$) than in a negative frame ($M = 3.32$, $SD = 1.47$). The main effect of the statistical format was insignificant, $F(1, 140) = 0.81$, $p =.776$, η_p^2 <.01.

More importantly, we found three important interactions. First, the interaction between statistical format and review quantity was significant, $F(1, 140) = 153.06$, $p <.001$, η_p^2 =.52; see Figure 2a, left side). When the total number of reviews for a product was large, participants exhibited a higher willingness to buy in the percentage format ($M= 4.32$, $SD = 1.62$) than in the frequency format ($M = 3.88$, $SD = 1.60$), $F(1, 140) = 62.15$, $p <.001$, η_p^2 =.31. Conversely, when the product had a small number of total reviews, participants showed a higher willingness to buy in the frequency format ($M= 3.97$, $SD = 1.45$) than in the percentage format ($M= 3.50$, $SD = 1.44$), $F(1, 140) = 59.91$, $p <.001$, η_p^2 =.30, supporting H1.

Second, the interaction between the statistical format and review valence was significant, $F(1, 140) = 130.83$, $p <.001$, η_p^2 =.48; see Figure 2b, left side). When the review valence was positive, participants' purchase intentions were higher in the percentage format ($M= 4.75$, $SD = 1.32$) than in the frequency format ($M = 4.42$, $SD = 1.44$), $F(1, 140) = 40.30$, $p <.001$, η_p^2 =.22. However, when the review valence was negative, participants' purchase intentions were higher in the frequency format ($M = 3.42$, $SD = 1.436$) than in the percentage format ($M= 3.08$, $SD = 1.38$), $F(1, 140) = 36.71$, $p<.001$, η_p^2 =.21, supporting H2.

Third, the interaction between statistical format and review frame was significant, $F(1, 140) = 54.59$, $p <.001$, η_p^2 =.28; see Figure 2c, left side). In the positive frame, participants had a higher willingness to buy in the frequency format ($M= 4.65$, $SD = 1.33$) than in the percentage format ($M= 4.37$, $SD = 1.47$), $F(1, 140) = 20.48$, $p <.001$, η_p^2 =.13. In contrast, in the negative frame, participants had a higher willingness to buy in the percentage format ($M = 3.45$, $SD = 1.57$) than in the frequency format ($M = 3.20$, $SD = 1.35$), $F(1,140) = 22.51$, $p <.001$, η_p^2 =.14. This evidence supports H3.

Additionally, the four-way interaction among statistical format, review quantity, review valence, and review frame was significant, $F(1, 140) = 30.51$, $p <.001$, η_p^2 =.179; see Figure 2d). When the review quantity was small and the review valence positive, the frequency format in a positive review frame induced the highest purchase intention, $F(3, 560) = 69.74$, $p<.001$, η_p^2 =.27 (frequency format in a positive frame: $M = 5.00$, $SD = 1.21$; percentage format in a positive frame: $M = 4.64$, $SD = 1.22$; frequency format in a negative frame: $M = 4.04$, $SD = 1.32$; percentage format in a negative frame: $M = 3.71$, $SD = 1.16$).

When the review quantity was large and the review valence positive, the percentage format in a positive review frame induced the highest purchase intention, $F(3, 560) = 48.62$, $p<.001$, η_p^2 =.21 (percentage format in a positive frame: $M = 5.62$, $SD = 0.91$; frequency format in a positive frame: $M = 5.27$, $SD = 1.13$; percentage format in a negative frame: $M = 5.02$, $SD = 1.17$; frequency format in a negative frame: $M = 3.38$, $SD = 1.27$).

When the review quantity was small and the review valence negative, the frequency format in a positive review frame induced the highest purchase intention, $F(3, 560) = 31.58$, $p<.001$, η_p^2 =.15 (frequency format in a positive

frame: $M = 4.12$, $SD = 1.16$; percentage format in a positive frame: $M = 3.38$, $SD = 1.22$; frequency format in a negative frame: $M = 2.73$, $SD = 1.11$; percentage format in a negative frame: $M = 2.28$, $SD = 1.06$).

When the review quantity was large and the review valence negative, the frequency format in a positive review frame induced the highest purchase intention, $F(3, 560) = 109.10$, $p<.001$, $\eta_p^2 = .37$ (frequency format in a positive frame: $M = 4.22$, $SD = 1.42$; percentage format in a positive frame: $M = 3.85$, $SD = 1.41$; percentage format in a negative frame: $M = 2.80$, $SD = 1.29$; frequency format in a negative frame: $M = 2.62$, $SD = 1.19$).
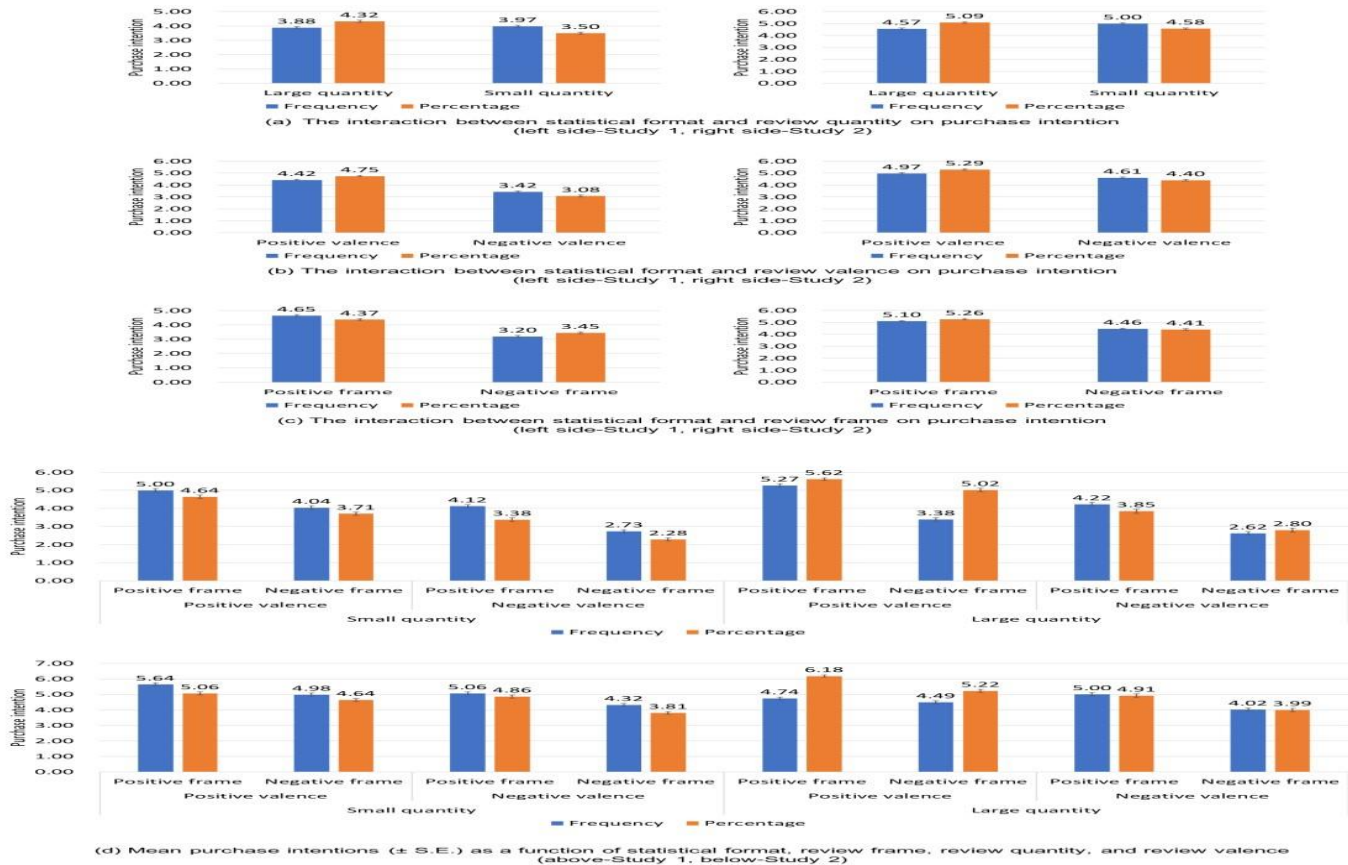

Figure 2

5   Study 2

5.1   Participants, Materials, and Procedure

In this study, the participants browsed through the products and online review information on the online shopping interface using a mobile terminal; however, we wanted to test our hypotheses in more realistic online shopping scenarios. Furthermore, to reduce mutual interference among experimental conditions, we adopted a between-subjects design: 2 statistical format × 2 review quantity × 2 review valence × 2 review frame. We also selected other products from the categories of clothing (a baseball cap), food (nuts), and electronic appliances (a television) to further test our hypotheses.

The variable manipulation was the same as in Study 1 (Table 3), but the experimental scenarios were more similar to real online shopping (Figure 3). The participants were asked to look at these online shopping interfaces and then to indicate their purchase intentions for the selected products (on a scale ranging from 1 = *very unwilling to buy* to 7 = *very willing to buy*). Cronbach's α =.69.

Table 3 *Experimental Materials (in Study 2)*

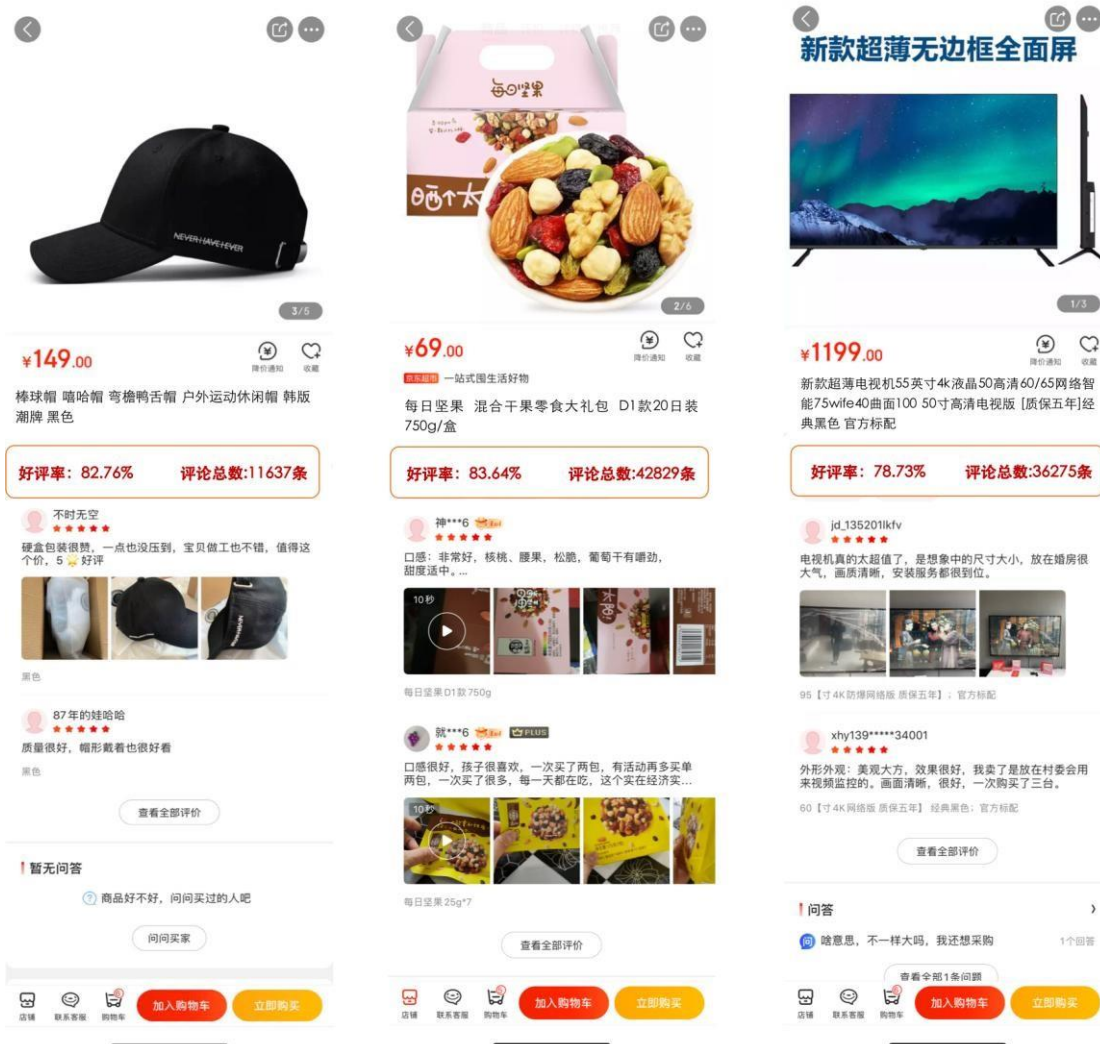|			Baseball cap		Nuts		Television	
			Frequency format	Percentage format	Frequency format	Percentage format	Frequency format	Percentage format
Negative review valence	Large review quantity	Positive frame	The number of positive reviews: 9,631; the number of reviews: 11,637	The percentage of positive reviews: 82.76%; the number of reviews: 11,637	The number of positive reviews: 35,822; the number of reviews: 42,829	The percentage of positive reviews: 83.64%; the number of reviews: 42,829	The number of positive reviews: 28,559; the number of reviews: 36,275	The percentage of positive reviews: 78.73%; the number of reviews: 36,275
		Negative frame	The number of negative reviews: 2,006; the number of reviews: 11,637	The percentage of negative reviews: 17.24%; the number of reviews: 11,637	The number of negative reviews: 7,007; the number of reviews: 42,829	The percentage of negative reviews: 16.36%; the number of reviews: 42,829	The number of negative reviews: 7,716; the number of reviews: 36,275	The percentage of negative reviews: 21.27%; the number of reviews: 36,275
	Small review quantity	Positive frame	The number of positive reviews: 96; the number of reviews: 116	The percentage of positive reviews: 82.76%; the number of reviews: 116	The number of positive reviews: 358; the number of reviews: 428	The percentage of positive reviews: 83.64%; the number of reviews: 428	The number of positive reviews: 285; the number of reviews: 362	The percentage of positive reviews: 78.73%; the number of reviews: 362
		Negative frame	The number of negative reviews: 20; the number of reviews: 116	The percentage of negative reviews: 17.24%; the number of reviews: 116	The number of negative reviews: 70; the number of reviews: 428	The percentage of negative reviews: 16.36%; the number of reviews: 428	The number of negative reviews: 77; the number of reviews: 362	The percentage of negative reviews: 21.27%; the number of reviews: 362
Positive review valence	Large review quantity	Positive frame	The number of positive reviews: 11,437; the number of reviews: 11,637	The percentage of positive reviews: 98.28%; the number of reviews: 11,637	The number of positive reviews: 42,431; the number of reviews: 42,829	The percentage of positive reviews: 99.07%; the number of reviews: 42,829	The number of positive reviews: 35,274; the number of reviews: 36,275	The percentage of positive reviews: 97.24%; the number of reviews: 36,275
		Negative frame	The number of negative reviews: 200; the number of reviews: 11,637	The percentage of negative reviews: 1.72%; the number of reviews: 11,637	The number of negative reviews: 398; the number of reviews: 42,829	The percentage of negative reviews: 0.93%; the number of reviews: 42,829	The number of negative reviews: 1,001; the number of reviews: 36,275	The percentage of negative reviews: 2.76%; the number of reviews: 36,275
	Small review quantity	Positive frame	The number of positive reviews: 114; the number of reviews: 116	The percentage of positive reviews: 98.28%; the number of reviews: 116	The number of positive reviews: 424; the number of reviews: 428	The percentage of positive reviews: 99.07%; the number of reviews: 428	The number of positive reviews: 352; the number of reviews: 362	The percentage of positive reviews: 97.24%; the number of reviews: 362
		Negative frame	The number of negative reviews: 2; the number of reviews: 116	The percentage of negative reviews: 1.72%; the number of reviews: 116	The number of negative reviews: 4; the number of reviews: 428	The percentage of negative reviews: 0.93%; the number of reviews: 428	The number of negative reviews: 10; the number of reviews: 362	The percentage of negative reviews: 2.76%; the number of reviews: 362

Figure 3

Based on the calculated sample size given above, we collected 963 valid participants (540 females; M_{age} = 30.70, SD = 7.85) from Sojump (http://www.Wjx.cn) and paid each of them ¥5 (¥1 = $0.14).

The thresholds for a small/large review quantity and a positive/negative review were set the same as in Study 1, and the post-check is reported in Table 2, suggesting a successful manipulation of review quantity and valence.

5.2   Results

A 2 statistical format × 2 review quantity × 2 review valence × 2 review frame ANOVA was conducted. The results showed that the main effect of review valence was significant, $F(15, 947) = 85.21$, $p < .001$, $\eta_p^2 = .08$. When the review valence was positive, the participants were more likely to buy the product ($M_{positive}$ = 5.13, SD = 1.14, $M_{negative}$ = 4.50, SD = 1.17). The main effect of review frame was significant, $F(15, 947) = 123.09$, $p < .001$, $\eta_p^2 = .12$: participants' purchase intentions were higher in a positive frame (M = 5.19, SD = 1.11) than in a negative frame (M = 4.43, SD = 1.16). The main effects of review quantity and statistical formatwere insignificant ($F(15, 947) = 0.123$, $p = .726$, $\eta_p^2 < .01$; $F(15, 947) = 0.619$, $p = .431$, $\eta_p^2 = .01$).

Similar to Study 1, we found a significant interaction between statistical format and review quantity, $F(15, 947) = 46.82$, $p < .001$, $\eta_p^2 = .05$; see Figure 2a–right side). When the total number of reviews was large, the participants exhibited a higher willingness to buy in the percentage format (M = 5.09, SD = 1.21) than in the frequency format (M = 4.57, SD = 1.16), $F(1, 947) = 29.11$, $p < .001$, $\eta_p^2 = .03$. Conversely, when the total number ofreviews was

small, the participants showed a higher willingness to buy in the frequency format (M= 5.00, SD = 1.12) than in the percentage format (M = 4.58, SD = 1.20), $F(1, 947)$ = 18.33, p <.001, η_p^2 =.02, supporting H1.

Further, the interaction between statistical format and review valence was significant, $F(15, 947)$ = 14.86, p <.001, η_p^2 =.02; see Figure 2b, right side). When the review valence was positive, participants' purchase intentions were higher in the percentage format (M= 5.29, SD = 1.09) than in the frequency format (M = 4.97, SD = 1.16), $F(1, 947)$ = 10.77, p =.001, η_p^2 =.01.

However, when the review valence was negative, participants' purchase intentions were higher in the frequency format (M = 4.61, SD = 1.14) than in the percentage format (M= 4.40, SD = 1.20), $F(1, 947)$ = 4.71, p =.030, η_p^2 =.01, supporting H2. It is noteworthy that we did not find a significant interaction between statistical format and review frame, $F(15, 947)$ = 1.81, p =.179, η_p^2 <.01, and, thus, Study 2 did not support H3. We discuss this in the General Discussion section.

Additionally, the four-way interaction among statistical format, review quantity, review valence, and review framewas significant, $F(15, 947)$ = 6.07, p =.014, η_p^2 =.006; see Figure 2d). Consistent with Study 1, when the number of reviews was small and the review valence positive, the frequency format in a positive review frame induced the highest purchase intention, $F(1, 947)$ = 15.93, p<.001, η_p^2 =.02 (frequency format in a positive frame: M = 5.64, SD = 0.86; percentage format in a positive frame: M = 5.06, SD = 1.28; frequency format in a negative frame: M = 4.98, SD = 1.03; percentage format in a negative frame: M =4.64, SD = 0.96).

When the review quantity was large and the review valence positive, the percentage format in a positive review frame induced the highest purchase intention, $F(1, 947)$ = 20.11, p<.001, η_p^2 =.02 (percentage format in a positive frame: M = 6.18, SD = 0.52; percentage format in a negative frame: M = 5.22, SD = 0.87; frequency format in a positive frame: M = 4.74, SD = 1.12; frequency format in a negative frame: M = 4.49, SD = 1.28).

When the review quantity was small and the review valence negative, the frequency format in a positive review frame induced the highest purchase intention, $F(1, 947)$ = 44.31, p<.001, η_p^2 =.05 (frequency format in a positive frame: M = 5.06, SD = 1.00; percentage format in a positive frame: M = 4.86, SD = 1.18; frequency format in a negative frame: M = 4.32, SD = 1.18; percentage format in a negative frame: M = 3.81, SD = 0.99).

When the review quantity was large and the review valence negative, the frequency format in a positive review frame induced the highest purchase intention, $F(1, 947)$ = 49.82, p<.001, η_p^2 =.05 (frequency format in a positive frame: M = 5.00, SD = 0.88; percentage format in a positive frame: M = 4.91, SD = 1.10; frequency format in a negative frame: M = 4.02, SD = 1.15; percentage format in a negative frame: M = 3.99, SD = 1.09).

6   General Discussion

This research explored whether the recently proposed "*love of large numbers*" theory always exists and whether consumers have biases in the processing of online review information in online shopping under the conditions of different statistical formats, review quantities, valences, and frames. The two main findings from our studies were as follows: 1) the frequency format induced higher purchase intentions than did the percentage format, with a small review quantity, a negative review valence, or a positive review frame; and 2) the percentage format induced higher purchase intentions than did the frequency format, with a large review quantity, a positive review valence, or a negative review frame.

Theoretically, these findings suggest that people are irrational in terms of processing information and that their purchase intentions are influenced by the way in which information is presented. This supports the theories of bounded rational decision-making, heuristics, and fuzzy-trace theory (Gigerenzer & Hoffrage, 1995; Reyna, 2004). The focus of information and the processing accuracy differed between the two statistical formats: the numerators were more focused in the frequency format, whereas the denominators (total reviews) were more

focused in the percentage format, and the process of ratio was less accurate in the frequency (vs. percentage) format. These differences led us to conclude that, with a positive review frame, a small review quantity, or a negative review valence, the frequency format was superior to the percentage format; and, with a negative review frame, a large review quantity, or a positive review valence, the percentage format was superior to the frequency format. However, it is worth noting that we did not observe a significant interaction effect between the statistical format and review frame in Study 2. We speculate that the emulation scenarios in Study 2 involved excess information (e.g., product price, titles and photos, reviews). This information may have diverted participants' attention to the numerators of the reviews, weakening the effect of the statistical format. More importantly, our results cast doubt on the finding of the *love of large numbers* (Powell et al., 2017). We found that it is not necessarily the total number of reviews that has a decisive effect on consumers' purchase intention; rather, we found an interaction between the number of positive (or negative) reviews and the format in which they were presented. For example, we found that consumers preferred the products with a large number of reviews only when the number of positive (or negative) reviews was presented in the percentage format, while the *love of large numbers* was not supported when the number of positive (or negative) reviews was presented in the frequency format.

Regarding practical applications, the current research provides inspiration and guidelines for online shopping platforms to design their presentation of products' online reviews. Our findings suggest that with respect to the frequency vs. the percentage formats, one is not always superior to the other; rather, there are variations according to review frame, quantity, and valence.

Our research explored the number of reviews from a multi-dimensional perspective and found that single presentation format of online reviews used by current platforms may result in consumers' perceptual bias. Therefore, the platform should present multi-dimensional information about the number of reviews in a standard way. For example, it should present not only the number of positive or negative reviews, but also the percentage in addition to the absolute number, so as to help consumers make more accurate and rational decisions.

6.1 Limitations

First, to ensure information equivalence, we presented total reviews in both statistical formats. In daily online shopping, online shopping platforms sometimes present percentage information without total reviews (e.g., positive review ratio: 96.41%). On the one hand, we speculate that, compared to the percentage format with total reviews, the percentage without total reviews may induce higher purchase intentions in cases in which there are fewer reviews because the disadvantage of total reviews is hidden. On the other hand, the percentage without total reviews may induce lower purchase intentions in cases with a large number of reviews because the advantage of total reviews is also hidden. Future studies should explore this topic further.

Second, in the manipulation of review valence, this study did not include the situation in which review valence is too negative (e.g., a less than 50% positive review ratio) because on real online shopping platforms, product review valences are generally between 80% and 100% (positive review ratio), and products with a review valence lower than 80% will be removed from online —shelves‖ (Ma et al., 2017). Future studies can further explore the effect of the statistical format in cases with a negative review valence.

References

Akl, E. A., Oxman, A. D., Herrin, J., Vist, G. E., Terrenato, I., Sperati, F., & Schünemann, H. (2011). Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database of Systematic Reviews*, (3):10.1002/14651858.cd006776.pub2

Brase, G. L. (2002). Ecological and evolutionary validity: Comments on Johnson-Laird, Legrenzi, Girotto, Legrenzi, and Caverni's (1999) mental- model theory of extensional reasoning uncertainty. *Psychological Review*, *109*, 722−728.

Chen, Y. F. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior*, *24* , 1977–1992.

China Internet network information center (CNNIC). (2016). 2015 Online Shopping Market Research Report in China

China Internet network information center (CNNIC). (2020). The 44th Statistical Report on Internet Development in China

Dackermann, T., Kroemer, L., Nuerk, H. C., Moeller, K., & Huber, S. (2018). Influences of presentation format and task instruction on children's number line estimation. *Cognitive Development*, *47*, 5362:10.1016/j.cogdev.2018.03.001

De Pelsmacker, P., Van Tilburg, S., & Holthof, C. (2018). Digital marketing strategies, online reviews and hotel performance. *International Journal of Hospitality Management*, *72*, 47-55:10.1016/j.ijhm.2018.01.003

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, *102*(4), 684-704:10.1037/0033-295x.102.4.684.

Hangzhou.gov.cn (2020, November 12). China in motion, Tmall Double 11 Witnesses Global "Move". Retrieved from http://www.hangzhou.gov.cn/art/2020/11/12/art_812266_59017125.html

Kahneman, D. and Tversky, A. (1979) Prospect Theory: An Analysis of Decision under Risk. Econometrica: Journal of the Econometric Society, 47, 263-291. http://dx.doi.org/10.2307/1914185 Katz, E., and Lazarsfeld, P.F.. *Personal Influence*. New Brunswick, NJ: Transaction Publishers, 2006.

Lee, B., Liu, J., Choung, H., & McLeod, D. (2019). Framing risk with numbers: The framing effects of risk assertions and number formats on emotions and risk perceptions. *Mass Communication and Society*, *22*(3), 344-364:10.1080/15205436.2018.1536790

Lu, X., Ba, S., Huang, L., & Feng, Y. (2013). Promotional marketing or word-of-mouth? Evidence from online restaurant reviews. *Information Systems Research*, *24*(3), 596-612:10.1287/isre.1120.0454

Ma, H., Zhao, F., & Yao Q. (2017). Ratio Bias in Online Shopping: Effect of Positive Comments and Involvement on Willingness to Buy. *Journal of Psychological Science*, 40(2), 436-441

McKechnie, S., Devlin, J., Ennew, C., & Smith, A. (2012). Effects of discount framing in comparative price advertising. *European Journal of Marketing*, *46*(11/12), 1501-1522:10.1108/03090561211259952

Nguyen, D. H., de Leeuw, S., & Dullaert, W. E. (2018). Consumer behaviour and order fulfilment in online retailing: a systematic review. *International Journal of Management Reviews*, *20*(2), 255-276:10.1111/ijmr.12129

Petrova, D., Joeris, A., Sánchez, M. J., Salamanca-Fernández, E., & Garcia-Retamero, R. (2018). How are risk ratios reported in orthopaedic surgery journals? A descriptive study of formats used to report absolute risks. *BMJ Open*, *8*(11), e025047:10.1136/bmjopen-2018-025047

Powell, D., Yu, J., DeWolf, M., & Holyoak, K. J. (2017). The love of large numbers: a popularity bias in consumer choice. *Psychological Science*, *28*(10), 1432-1442:10.1177/0956797617711291

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science*, 13, 60−66: 10.2307/20182911

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107.doi:10.1016/j.lindif.2007.03.011

Romero, M., W Craig, A., Kumar, A., & Mormann, M. (2018). ―4‖ Vs.―Four‖: The Influence of Number Format on Perceived Magnitude and Product Evaluations. *ACR European Advances*.

Shue, K., & Townsend, R. R. (2021). Can the Market Multiply and Divide? Non-Proportional Thinking in Financial Markets. *The Journal of Finance, 76*(5), 2307-2357: 10.1111/jofi.13059

Sotiriadis, M. D., & Van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, *13*(1), 103-124:10.1007/s10660-013-9108-1

Srivastava, J., & Koukova, N. (2018). Why Is 1 Out of 20 Riskier Than 5%? Effect of Representing Unlikely Events as Frequency Versus Percentage on Risk Perceptions. *ACR Asia-Pacific Advances*.

Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. (2006). Formats for improving risk communication in medical tradeoff decisions. *Journal of Health Communication*, *11*(2), 167182:10.1080/10810730500526695

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

Yang, J., Sarathy, R., & Lee, J. (2016). The effect of product review balance and volume on online Shoppers' risk perception and purchase intention. *Decision Support Systems*, *89*, 66-76:10.1016/j.dss.2016.06.009

Yin, D., Bond, S., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *Mis Quarterly*, *38*(2), 539-560.

Zhang, Z., Zhang, Z., Wang, F., Law, R., & Li, D. (2013). Factors influencing the effectiveness of online group buying in the restaurant industry. *International Journal of Hospitality Management*, *35*, 237245:10.1016/j.ijhm.2013.06.012

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133-148:10.1509/jm.74.2.133