Two Essays on the Interplay Between Online Reviews and Consumers

By

Seyedjavad Mousavi
University of Kansas, 2022
MBA, University of Tehran, Tehran, 2017
B.Sc., Sharif University of Technology, Tehran, 2014

Submitted to the graduate degree program in the School of Business and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____

Chair: Dr. Surendra N. Singh

____

Dr. Promothesh Chatterjee

____

Dr. Prakash Shenoy

____

Dr. Sanjay Mishra

____

Dr. Steve Hillmer

Date Defended: June 17th, 2022

The dissertation committee for Seyedjavad Mousavi certifies that this is the approved version of the following dissertation:

Two Essays on the Interplay Between Online Reviews and Consumers

Chair: Dr. Surendra N. Singh

Date Approved: June 17th, 2022

Abstract

Online reviews and ratings have evolved into valuable resources for consumers as well as businesses. Given that an astronomical number of reviews are being posted on online review platforms, the roles of consumers and review platforms need to be investigated. This dissertation, therefore, investigates the interplay between online reviews and consumers. In Essay 1, the impact of various graphical representations of online rating distributions on consumer decision-making is studied. While displaying rating distributions in proportional versus simple bar graphs leads to lower product evaluations, it reduces other cognitive biases. This negative impact is reduced with an increase in the distribution's peak value. Essay 2 investigates how consumers' characteristics affect the textual content of negative online reviews. Consumers write negative reviews to complain about various aspects of products and services. We find that compared with liberal individuals, conservative consumers are more likely to complain about their dissatisfactory personal interactions with service employees in their reviews. As well as making significant contributions to both practical and theoretical issues, the two essays offer valuable ideas for future research.

Table of Contents

Chapter 1: Essay I—Unveiling Stars: How Graphical Displays of Online Consumer Ratings Affect Consumer Judgment and Choice

Unveiling Stars: How Graphical Displays of Online Consumer Ratings Affect Consumer Judgment and Choice

Abstract

Prior research has indicated that online consumer reviews significantly influence consumers' decisions. However, existing research has focused mainly on online review attributes (e.g., average ratings) that firms do not fully control; only limited research has investigated how controllable attributes (e.g., review display formats) affect consumers. Drawing on the research on graphical perception and proportion judgments, the authors examine the effectiveness of two prominent graphical display formats used by major e-commerce platforms—one that displays rating distributions in a proportional format (e.g., Amazon) and one that does so in a simple format (e.g., Google). By implementing a multimethod approach, the authors find that consumers respond less favorably to an item when its rating distribution is displayed in a graphically proportional format. Furthermore, this effect is attenuated as the distribution's peak value increases. The authors discuss the contributions to the marketing literature and provide insights that can enable managers to make more informed decisions.

Keywords: online consumer reviews, rating distribution, proportion judgment, graphical displays

Introduction

With the burgeoning of e-commerce, just about everything is reviewed on third-party platforms and e-commerce websites, offering consumers access to a vast array of online word of mouth (WOM). Given the enormous number of posts, it is impossible for visitors to these websites to read all the reviews that a product or service receives. Review platforms provide meaningful summary information (e.g., products' average rating, the number of reviews, and rating distributions) extracted from online consumer reviews to facilitate decision-making. While most platforms offer the same types of summary information, they typically use a wide variety of visual formats to exhibit the distribution of individual ratings (e.g., bar graphs in different forms, orientations, colors, and sizes) and their weighted averages (e.g., stars, bubbles, and squares).

Prior studies have extensively investigated the impact of the various attributes of online WOM, including the textual content of reviews, the number of reviews, average ratings, and their variance, on consumer decision-making (e.g., Chevalier and Mayzlin 2006; Rocklage and Fazio 2020; Rosario Babić et al. 2016; Schoenmueller, Netzer, and Stahl 2020; Wang, Liu, and Fang 2015; Watson, Ghosh, and Trusov 2018). However, the research examining how the visual representation of review summary information affects consumers' purchase decisions is limited (Fisher, Newman, and Dhar 2018; Rozenkrants, Wheeler, and Shiv 2017). The primary question we address in the current research is whether identical online rating distributions displayed in various formats can have different influences on consumer evaluation and choice. More specifically, we examine the impact of two classes of distribution formats widely used by prominent platforms on consumer judgment―proportional bar graphs (used by Amazon, Trustpilot, and TripAdvisor; Figure 1: left panel) and simple bar graphs (used by Google, Yelp, and Facebook; Figure 1: right panel).

[INSERT FIGURE 1 ABOUT HERE]

This aspect of online WOM is crucial because (1) previous research has shown that consumers rely heavily on rating distributions in their decision-making (Fisher et al. 2018; He and Bond 2015; Rozenkrants et al. 2017), (2) surveys have shown that the rating distribution summary is the most utilized feature of online WOM by consumers (Baymard 2017), and (3) unlike many other attributes, the presentation aspects of online WOM summary information are fully controlled by review platforms.

As Figure 1 shows, the fundamental difference between a proportional and simple bar graph is that the former (left panel in Figure 1) graphically displays the value of each star rating as a part-to-whole relationship, whereas in the latter (right panel in Figure 1), each star rating bar offers only information about its own value. Therefore, a proportional bar graph provides two fixed reference points—0 and all (100%) of the ratings—against which each bar can be visually compared. In other words, the value of each star rating in a proportional bar graph is shown using two areas. In one area, the proportion of people who have given a specific rating score (e.g., 5 stars) is highlighted. The other area, which is either lightly shaded or is the blank space between the end of the bar and the explicitly written values, represents those who have not assigned that rating score to the item, making part-to-whole relationships graphically available. In contrast, a simple bar graph displays information as individual values rather than as parts of a greater value by highlighting only the group of people who provided a specific rating, which offers only one fixed reference point (i.e., 0).

Drawing on the research on graphical perception and proportion judgment (e.g., Bonato et al. 2007; Hollands and Dyre 2000; Hollands and Spence 1998, 1992; Varey, Mellers, and

Birnbaum 1990), we contend that consumers evaluate an item less favorably when its rating distribution is displayed in a proportional rather than simple format. Specifically, we draw attention to a unique aspect of proportional bar graphs: graphical part-to-whole relationships. We argue that consumers rely on these graphical part-to-whole relationships to make sense of rating distributions and assess target items. Previous research has found that individuals tend to overestimate small proportions and underestimate large proportions (Hollands and Dyre 2000; Varey et al. 1990). Given that a copious amount of research (e.g., Hu, Pavlou, and Zhang 2017; Moe, Netzer, and Schweidel 2017; Schoenmueller, Netzer, and Stahl 2020) has indicated a positive imbalance in reviews (i.e., an overwhelming majority of reviews average four and above on a 5-point scale), we contend that the graphical presence of part-to-whole relationships dampens the evaluation in a proportional format vis-à-vis a simple format. For instance, we argue that consumers' subjective values of 5- and 4-star ratings (53% and 34%, respectively) in a proportional bar graph (as exhibited in the top panels of Figure 1) are lower than their subjective values in a simple bar graph because of consumers' tendencies to underestimate large proportions. We present five multimethod studies in the main manuscript and four in the Appendices in support of our proposition.

This research has important implications from both theoretical and managerial perspectives. First, this work extends the existing literature by studying a neglected feature of online WOM. Although prior research has suggested that presenting rating distributions in graphical displays affects consumers' evaluation of ratings (e.g., Fisher, Newman, and Dhar 2018; He and Bond 2015), it is still unclear how various graphical representations of ratings may affect consumers' interpretation of them. The current research scrutinizes the visualization aspect of online ratings to provide valuable insights into how and why rating distributions presented in various graphical formats may engender different judgments and choices. Second, by identifying the graphical format of rating distributions as an essential factor in the formation of consumers' judgments of online ratings, we present online review platforms and businesses with recommendations that can improve their performance. Third, by demonstrating how different graphical representations of ratings affect consumers' perceptions of those ratings, the findings of the current research contribute to graphical perception research as well as to the prior research that has found systemic biases in the way individuals evaluate proportions.

In the ensuing sections, we discuss the relevant literature and develop our conceptual framework and specific hypotheses. We then report the findings of nine studies designed to test those hypotheses and the underlying processes. Finally, we discuss the theoretical and managerial implications of this work and some avenues for future research.

Conceptual Background

Displaying rating distributions, as opposed to statistics such as average ratings, can lead to very different outcomes (e.g., Fisher, Newman, and Dhar 2018; He and Bond 2015; Rozenkrants, Wheeler, and Shiv 2017). Prior research, however, has focused mainly on how consumers incorporate distribution summary information into their decisions and the situational factors that influence these decisions. For instance, research has found that consumers tend to engage with rating distributions via categorization, summation, and deduction processes rather than via the mere extraction of information (Fisher et al. 2018). Consumers use distribution graphs to infer the level of dispersion (or consensus) in ratings, which influences their evaluations (Khare, Labrecque, and Asare 2011). Moreover, the same rating distribution affects consumers' decisions and choices differently depending on the product category (He and Bond

2015) or situational factors such as the activation of self-expression goals (Rozenkrants et al. 2017).

Utility maximization models suggest that the distribution of a set of ratings should generate equivalent evaluations and choices, regardless of its presentation format (Tversky and Kahneman 1986). Empirical evidence (Cleveland, Diaconis, and McGill 1982; Kahneman and Tversky 1979; Stone, Yates, and Parker 1997), however, shows otherwise because varied representations of the same information highlight its different aspects and thus alter cognitive processes, influencing task performance (Bettman and Kakkar 1977; Kim and Lakshmanan 2021) and consumer preferences (Bettman, Luce, and Payne 1998). Advancing this line of reasoning, we predict that providing part-to-whole relationships graphically by displaying rating distributions in a proportional (vs. simple) format affects consumers' cognitive processes and outcomes. Moreover, we argue that the graphical representations of part-to-whole relationships propel consumers into distinct perceptual and cognitive tasks, which lead to lower evaluations. For deeper insight into the underlying processes, we first examine the two graphical formats through the lens of length and proportion judgments and develop our hypotheses. Then, we discuss other differences that may lead to the same or other hypotheses.

Biases in Length and Proportion Judgments

Research on graphical displays has found that consumers tend to give more weight to graphical information than to verbal information (Lurie and Mason 2007). Therefore, in both simple and proportional bar graphs, consumers' decisions should be more influenced by graphical cues (i.e., bars) than by textual information (i.e., written frequencies or percentages). However, the fundamental differences in graphical cues provided by each format lead to distinct cognitive tasks. Prior research has found that while perceptions of changes and differences are direct with simple bar graphs (i.e., they require fewer mental operations), perceptions of proportions are direct with proportional bar graphs (Hollands and Spence 1992). Therefore, while simple bar graphs naturally lead consumers to compare the bars through length judgments, individuals process proportional bar graphs by engaging in proportion judgments. As a result, to examine how consumers may be influenced by each graphical format, we should first discuss how length and proportion judgments are made and what factors may influence these judgments.

Length judgments. Previous research has shown that individuals are quite accurate in judging lengths (Cleveland and McGill 1986, 1987). For example, Cleveland and McGill (1984) find that judging and comparing the length of graphical cues is among the easiest perceptual tasks that can be performed by an individual. On the other hand, Hollands and Dyre (2000) report a perfectly linear relationship between the true length of a stimulus and individuals' subjective judgment of that length. In a simple bar graph, the only informative attribute of a bar is its length. Therefore, given that (a) a linear relationship exists between the objective and subjective lengths of a stimulus, and (b) the relative lengths of bars are determined by their values, engaging in length judgments should not lead to a biased perception of ratings.

Proportion judgments. Unlike length judgments, research has identified two biases in the judgment of proportions: small proportions are overestimated, while large proportions are underestimated (Gonzalez and Wu 1999; Hollands and Spence 1992; Hollands, Tanaka, and Dyre 2002; Landy, Guay, and Marghetis 2018; Slusser and Barth 2017; Slusser, Santiago, and Barth 2013; Tversky and Kahneman 1992; Zhang and Maloney 2012; Zhang, Ren, and Maloney 2020). Proportion judgments have been studied mainly in the context of perceptual magnitude judgments, where biased assessments of part and whole magnitudes lead to biased estimates of proportions (e.g., Hollands and Dyre 2000; Hollands, Tanaka, and Dyre 2002; Sullivan et al. 2011). A consistent finding of proportion judgment research is that individuals tend to overestimate small proportions and underestimate large proportions (i.e., an inverted S-shaped curve), which suggests that when judging proportions, the subjective perception of a large [small] proportion is smaller [greater] than its true value (Figure 2).

[INSERT FIGURE 2 ABOUT HERE]

This bias pattern, however, is not limited to graphical representations and has been observed in various contexts, including gambling and probability judgments, ratio and proportion judgments, time perception, and scale reading (for a review, see Holland and Dyre 2000). Tversky and Kahneman's (1992) cumulative prospect theory has also documented the same pattern in proportion judgments. They argue that:

> *Overweighting of small probabilities contributes to the popularity of both lotteries and insurance. Underweighting of high probabilities contributes both to the prevalence of risk aversion in choices between probable gains and sure things, and to the prevalence of risk seeking in choices between probable and sure losses*. (p. 316)

A crucial finding of this research is the extent to which individuals’ judgments of proportions are biased. As displayed in Figure 2, the estimation errors (i.e., the subjective proportion minus the true proportion) committed by individuals when judging small proportions are much smaller than the estimation errors for large proportions. Consequently, we expect the underestimation of large proportions to be more influential in determining how consumers evaluate rating distributions displayed in proportional bar graphs. Consistent with these findings, we predict that displaying rating distributions in proportional bar graphs lowers consumers’ judgment of large proportions (i.e., positive ratings) and, consequently, their judgment of the target item. To explain how those underestimations may influence consumers, we next discuss the findings of previous research on how consumers process rating distributions in different ways and how these processes may be influenced by proportion and length judgments.

Consumers' Processing of Rating Distribution and Hypothesis Development

In both simple and proportional bar graphs, consumers extract bar values and utilize this information in different ways (Fisher et al. 2018; Rozenkrants et al. 2017; West and Broniarczyk 1998). In this section, we discuss how graphical proportions may influence the tasks that are involved in processing the distribution of a set of ratings and how the effect of the graphical format is influenced by the shape of the rating distribution.

Information extraction. Extracting values is the first and most basic task in processing a rating distribution. While consumers extract information from simple bar graphs by assessing bar lengths, the extraction of information from proportional bar graphs is influenced by automatic proportion judgments. Given that an overwhelming majority of reviews average four and above on a 5-point scale (Hu et al. 2017; Moe et al. 2017; Schoenmueller et al. 2020), 4- and 5-star ratings account for larger proportions of all ratings in most cases, and as a result, 1- and 2-star ratings account for smaller proportions. This finding is also confirmed by our analysis, which revealed that 86% [83%] of the reviews posted on Google [TripAdvisor] in Study 1 had either four or five stars. Therefore, a lower subjective judgment of large proportions with a proportional bar graph should lead to an underestimation of positive ratings. Hence, we expect consumers to form a lower evaluation of an item when its rating distribution is displayed in a proportional (vs. simple) bar graph. Formally, we propose the following:

H_1: *Since the subjective values of large proportions are smaller than their objective values, using a proportional (vs. simple) bar graph leads to a lower evaluation of the target item.*

Role of peak value. A distribution's peak value (i.e., the value of the tallest bar) can range between 20% (when ratings are distributed uniformly) and 100% (when any single bar has all ratings). Peak value, therefore, is a useful indicator of how an item's ratings are distributed. Distributions with high peak values include fewer large proportions than do those with smaller peak values. For instance, while a distribution whose peak value is 90% has only one large proportion (i.e., the peak value itself), a distribution with a peak value of 40% is likely to have multiple large proportions. As suggested in Figure 2, the difference between the subjective judgment and the true value of a proportion is smaller at the two ends or boundaries (i.e., estimation errors are smaller for proportions closer to 0 or 1). One explanation for this phenomenon is that there is little room for error at the two boundaries (Hollands and Dyre 2000; Spence and Krizel 1994). Therefore, with an increase in the peak value, the underestimation of large proportions diminishes for two reasons: (1) fewer large proportions, which can be underestimated, exist, and (2) underestimation sizes are smaller for proportions that are closer to one (i.e., a large peak value). Consequently, the negative effect of using proportional bar graphs, as described in H_1, should be moderated by the peak value. In other words, we propose the following:

H_2: *As the peak value increases, the negative effect of displaying rating distributions in the proportional format becomes weaker, leading to a positive, two-way interaction between the graphical format and the peak value of the distribution.*

Categorical thinking and binary bias. Information extraction is the first step in evaluating rating distributions. Aggregating 4- and 5-star [1- and 2-star] ratings and perceiving them as being equally positive [negative] is another crucial task undertaken by consumers to process rating distributions, resulting in a binary assessment of ratings (Fisher et al. 2018). This process, therefore, leads to the imbalance score (defined as the difference between the proportion of 5- and 4-star reviews and that of 1- and 2-star reviews) having a positive impact on product evaluations, even after controlling for other statistics such as average rating, peak value, and variance of ratings.

We propose that the graphical format in which ratings are displayed should influence such mental operations in at least two ways. First, given that perceptions of changes and differences are direct with simple bar graphs, consumers who view simple (vs. proportional) bar graphs should be more likely to carry out mental calculations found in binary bias research. Second, a lower subjective value of positive ratings (as discussed previously) in proportional (vs. simple) bar graphs should lead to a lower subjective value of the imbalance score. Therefore, we expect binary bias to be less influential when rating distributions are displayed in a proportional (vs. simple) bar graph. More formally, we propose the following:

H_3: The positive impact of the imbalance score on product evaluation is weaker when distributions are displayed in the proportional (vs. simple) bar graph format.

In addition to the presence or lack of graphical proportionalities, simple and proportional bar graphs differ in several other aspects, including bars' relative and absolute lengths, mutable and fixed reference points, and the number of graphical cues provided. These differences, therefore, may also lead to altered judgments and may make the processing of one format easier than that of the other. In the next section, we examine these differences and discuss the extent to which they may support or undermine our predictions through two lenses: visual perception and processing fluency.

Differences Beyond the Presence of Part-to-Whole Relationships

Absolute bar lengths. When the same space is dedicated to both graphs, because of the nature of the two formats, all bars in simple bar graphs appear taller and consequently more visually salient than those in proportional bar graphs. Therefore, one explanation for our main proposition that proportional bar graphs lead to lower evaluations (H_1) can be that this difference in the absolute length of the bars and their visual salience may alter the perception of ratings and items' favorability. In other words, consumers may evaluate simple bar graphs more positively simply because the bars in such graphs are taller and more salient. While this potential explanation seems plausible, at least one strong counterargument exists. Both negative and positive bars in a simple bar graph appear taller with the same ratios. Therefore, the underlying assumption of this explanation is that the effect of increasing the saliency of positive ratings is stronger than that of negative ratings (i.e., a positivity effect). However, the research on the negativity effect (Ahluwalia 2002; Klein and Ahluwalia 2005) suggests otherwise.

Relative length judgments. Another visual difference between simple and proportional bar graphs is how bars appear differently relative to bar graph borders (i.e., chart area). Thus, the second explanation for why proportional bar graphs lead to lower evaluations (H_1) could be that instead of proportional judgments, consumers judge the size of the bars relative to the chart area, and the smaller relative size of the bars in proportional bar graphs leads to a lower evaluation. This explanation is unlikely, as it also requires the positivity effect assumption mentioned earlier. We test both absolute and relative length judgment explanations in Study 3.

Number of graphical cues. The third possible explanation for why proportional graphs lead to lower evaluations (H_1) could be rooted in how visually crowded they are. A proportional bar graph has more graphical cues than does a simple bar graph (see Figure 1's top panels). The mere presence of those additional shaded rectangles may lead to lower evaluations. The crowdedness of proportional bar graphs may also lower their visual appeal, leading to lower evaluations. While these explanations seem plausible, and we test them in our studies, they do not predict any outcome regarding the moderating role of the peak value (H_2) and how binary bias (considering high rating scores equally positive and low scores equally negative) is influenced by the graphical format (H_3).

Graphical Formats and Processing Fluency

Processing fluency is defined as the subjective experience of the ease with which incoming information is processed (Lee and Labroo 2004). Research has shown that higher processing fluency can lead to a higher evaluation of the target stimulus (Shen, Jiang, and Adaval 2010), as consumers tend to misattribute the fluency of information processing to the stimulus itself and associate fluent stimuli with greater appeal (Schwarz 2004). Before discussing the differences in the two graphical formats (i.e., simple and proportional bar graphs), it should be noted that the processing fluency research suggests that consumers should evaluate both distribution graphs more positively or at least as positively as distribution tables because graphical perception research has shown that offering information in a graphical format facilitates cognitive processes (Lurie and Mason 2007; Scaife and Rogers 1996). However, it is not apparent whether simple or proportional bar graphs are easier to process, but this may become more evident when examined through the following vantage points.

Complexities incurred by the number of graphical cues and lack of information. Graphical perception research finds that the number of elements displayed in a visual stimulus is essential in determining its complexity: the more graphical cues a stimulus has, the more complicated it is (Hegarty 2011). According to this view, processing a distribution displayed in a proportional bar graph should be more complicated than when it is displayed in a simple bar graph. The lower processing fluency caused by a higher number of cues can lead to lower product evaluations and explain why the use of proportional graphs leads to lower evaluations of the target item (H_1). Conversely, a proportional bar graph offers more information than does a simple bar graph. Therefore, a lack of information can make the processing of a simple bar graph distribution more complex, leading to lower evaluations, which contradicts H_1.

Complexities incurred by shifts in reference points. As discussed earlier, the two formats lead to crucial differences in the reference points against which the values are compared. In a proportional bar graph, a bar's apparent length is determined solely by its own value, meaning that a full [half empty] bar always represents 100% [50%] of the ratings. Thus, the reference point against which the value of each bar is visualized (and therefore compared) is always a constant number (i.e., 100% of the ratings), leading to proportion judgments. In a simple bar graph, however, the length of a bar is determined by its value as well as the distribution's peak value. In simple bar graphs, therefore, the reference point against which individual bars are visualized is the distribution's peak value, which changes from one distribution to another. The crucial question, therefore, is to what extent do mutable reference points (i.e., peak values) affect the processing fluency of simple bar graphs and consumers' evaluations?

From this perspective, we contend that the processing fluency of simple bar graph distributions depends on the number of stimuli being processed at a time or in a sequence (i.e., the number of items being evaluated). When evaluating a single item, no initial reference point that needs to be changed exists. However, when multiple simple bar graph distributions are involved (e.g., in the sequential evaluation task in Studies 2 and 3), changing reference points from one distribution to another should make the processing of simple bar graph distributions more difficult. Thus, while processing one simple bar graph distribution should not be more complicated than processing its proportional bar graph counterpart, processing multiple simple bar graph distributions should be more complicated. Therefore, a processing fluency explanation predicts no difference in single-item evaluation tasks but predicts lower evaluations of simple bar graph distributions when multiple products or services are involved. In other words, this interpretation of processing fluency predicts that we should observe no difference in the evaluations of simple and proportional bar graph distributions in between-subjects studies (Appendices A, B, and C) and the opposite of the central proposition (H_1) in mixed-design studies (Studies 2 and 3 and Appendix D). Moreover, this interpretation predicts that consumers who view rating distributions in simple bar graphs require more time processing the information (Westerman, Lanska, and Olds 2015) than those who view proportional bar graphs (Studies 3 and 4 and Appendix A).

Overview of Studies

We report the results of nine studies, including five in the current manuscript and four in the Appendices. Study 1 is an observational study where we test whether the ratings that are provided for a particular business differ on two platforms that use different graphical formats. In Study 2, we ask online participants to evaluate several target items whose ratings are displayed in simple or proportional bar graphs using a large sample of 200 unique distributions. This study provides a robust test of the effect of the graphical format and its generalizability in a controlled setting, testing H_1 (proportional graphs lead to lower ratings than do simple bar graphs), H_2 (as peak values increase, the negative effect of proportional graph format lessens), and H_3 (the impact of the imbalance score is weaker in the proportional bar graph format). In Study 3, we use five graphical formats of different sizes and visual appearances along with a table format to test alternative explanations. Study 4 tests the roles played by proportion and length judgments in the perception of ratings, and finally, in Study 5, we examine the effect of the graphical format on consumers' choices. Appendices A, B, C, and D provide additional support for our propositions in different contexts.

In Studies 2 and 3, we administered Fisher, Newman, and Dhar's (2018) four-item, ten-point Likert scale ["How do you feel about this product/service (unfavorable/favorable)?" "How would you expect your experience of this product/service to be (very negative/very positive)?" "How likely would you be to buy this product/service (very unlikely/very likely)?" and "How much would you be willing to pay for this product/service (not a lot/very much)?"] in randomized order. Study 4 asked participants to estimate the number of ratings associated with a star-rating when ratings were displayed in simple or proportional bar graphs.

In each experiment, a between-subjects design was used for the graphical format, where at least two graphical formats were included. Moreover, each experiment involved a unique sample of participants who had not taken part in previous studies. For experiments performed via Amazon Mechanical Turk (MTurk) and Prolific users, we administered an attention test at the beginning, allowing only those who followed the instructions to participate in the experiments (Appendix E).

Study 1: Do Consumers Rate the Same Businesses on Various Platforms Differently?

This study aimed to test the extent to which the ratings of a product or service might differ across platforms, especially on those that display rating distributions in different formats.

This question is crucial because the format in which ratings are displayed may also affect future customer ratings. We searched Google Maps, which uses the simple bar graph format, for restaurants in New York City. The search yielded 429 locations whose rating distributions and price ranges were collected. We then searched all 429 locations on TripAdvisor—which uses the proportional bar graph format—and found matches for 343 restaurants. From TripAdvisor restaurants' pages, we collected the number of reviews and their distributions and whether those restaurants had received any Michelin stars.

To test whether businesses had received different ratings on the two platforms, we conducted a mixed-effects regression model using individual rating scores as the dependent variable and platform as the independent variable. We included price level, Michelin star status, and their interactions with the platform as covariates. We treated price level as a continuous variable spanning from one to four, depending on the number of $ signs displayed on each restaurant page on Google. A random intercept for each restaurant (SD = .19) was also included in the model.

Results. The negative main effect of the platform on ratings revealed that restaurants in our dataset had received lower ratings on TripAdvisor (b = –.18, SE = .01, $p < .001$). Michelin restaurants were rated more favorably than non-Michelin restaurants (b = .13, SE = .03, $p < .001$), but the negative interaction between platform and Michelin star (b = –.03, SE = .01, $p < .001$) revealed that the positive effect of having Michelin stars was smaller for TripAdvisor. The main effect of price level on ratings was not significant (b = .001, SE = .01, $p = .970$), but its interaction with the platform revealed that higher prices led to higher ratings on TripAdvisor (b = .01, SE = .003, $p < .001$; see Table 1).

[INSERT TABLE 1 ABOUT HERE]

We next tested whether other more crucial attributes affecting the distribution of a set of ratings were different on the two platforms. We calculated two peak values and two variances for each restaurant: one based on ratings on Google and the other based on those on TripAdvisor. We conducted a paired t-test to examine the effect of the platform on the peak value and found that Google users had rated restaurants more homogeneously (mean peak value = 65.31%) than had TripAdvisor users (mean peak value = 55.74%; $t(342) = 21.01$, $p < .001$). Moreover, while the mode of all 343 rating distributions on Google was 5, on TripAdvisor the modes varied as follows: 5 (292 distributions), 4 (48 distributions), and 2 (1 distribution), with two distributions being multimodal.

A paired t-test also showed that there was a significant difference in imbalance scores—defined as the difference between the proportion of positive (4- and 5-star) and negative (1- and 2-star) ratings—between the two platforms ($M_{Google} = 79.67\%$, $M_{TripAdvisor} = 75.69\%$, $t(342) = 7.76$, $p < .001$). The *variance of all ratings* provided for restaurants in our sample was not significantly different on the two platforms ($F(608346, 356122) = .98$, $p = 1$). To test whether the variance of the ratings *for each restaurant* was different on the two platforms, we conducted an equality of variance test (F-test) for each restaurant, comparing the variance in their ratings on Google and TripAdvisor. Out of the 343 F-tests, we found that at least 126 restaurants (36.73% of our sample) had different variances on these platforms (*ps* < .1). The averages of the rating variances for these restaurants on Google and TripAdvisor were 1.06 and .80, respectively ($t(125) = 15.96$, $p < .001$), suggesting a lower consensus among Google users who rated those restaurants.

These results suggest that different sets of ratings are displayed for the same businesses or products on various platforms, for which there can be several justifications. First, Google and

TripAdvisor users may systematically differ in terms of several attributes, such as age, gender, education, expertise, and income, which can lead to different rating behaviors. Second, different perceptions of the same reviews and ratings can also influence consumers' expectations and thus the manner in which they rate an experience. Third, and more importantly, even when consumers have the same expectations and experiences, the positivity or negativity of prior ratings can influence how they rate products and businesses (Moe and Schweidel 2012; Schlosser 2005). For example, being exposed to negative opinions about a product may activate reviewers' concern regarding the social outcomes of their evaluations (e.g., being judged as having a low standard in case of positive evaluations), therefore leading them to adjust their evaluations downward and give the product lower ratings (Schlosser 2005).

This phenomenon may partly explain why TripAdvisor users had given lower ratings to restaurants than had Google users; displaying ratings using proportional bar graphs could negatively influence users' perception of the current ratings, trigger impression management goals among them, and lead them to provide lower rating scores. While the sources of these differences are not within the scope of the current work, such differences make studying the effects of using different graphical formats even more crucial. We suggest that different representations of ratings can indeed influence the future ratings of products/services. The following study examines how using simple and proportional bar graphs affects consumers' evaluations.

Study 2: Testing the Effect of the Graphical Format on Product Evaluation

Study 2 aimed to test the effect of the graphical format on consumer evaluation (H_1) in a controlled setting. The additional objectives of this study were to test the interaction effects between the graphical format and peak value and imbalance score described in H_2 and H_3. Per our conceptualization, the peak value of distribution should influence consumers' perception of proportions provided in a proportional bar graph (H_2). We contend that in a distribution with a low peak value (vs. a high peak value), the presence of part-to-whole relationships leads to a greater underestimation of positive ratings and further dampens consumers' evaluations. In addition, this study tested whether the graphical presence of part-to-whole relationships in the proportional bar graph mitigated the positive effect of the imbalance score (H_3).

Method. We recruited 1,192 participants (44% female, mean age = 37.29 years) from MTurk in return for a small payment. This study involved a 2 (graphical format: simple vs. proportional bar graph, as displayed in the top panels of Figure 1) × 2 (product: laptop vs. stapler) × 10 (rating distributions) mixed design. The first two factors varied between subjects, and the third varied within-subjects. Each participant saw ten randomly selected distribution graphs in a randomized order and was told that each distribution graph displayed a product's ratings. We measured product evaluation using the four-item scale mentioned earlier (Fisher et al. 2018). Since previous research suggests that any conclusion about graphical displays of information should be based on a large number of stimuli (Hegarty 2011), our distribution pool consisted of 200 randomly selected distributions from 130,000 randomly generated distributions (see Appendix F), resulting in 400 stimuli. Participants saw only a rating distribution displayed in a simple or proportional graph for each product evaluation task, as exhibited in the top panels of Figure 1. We used staplers and laptops as target products to test the effect of the graphical format on both small and relatively more significant purchases.

Results. We used a linear mixed-effects regression model with product evaluation as the outcome variable and graphical format (–1 = simple bar graph, +1 = proportional bar graph; H_1), average rating, variance, average rating × variance, distribution's peak value, graphical format × distribution's peak value (H_2), imbalance score and its interaction with the graphical format (H_3), product, and its interaction with the graphical format as fixed effects. We included variance in our model because research has found its negative impact on consumers' evaluations (Sun 2012). Moreover, because previous research has found that consumers show lower purchase intentions for products that do not have any negative ratings (i.e., low variance at higher rating averages; Maslowska, Malthouse, and Bernritter 2017), average rating × rating variance was added with the expectation of it having a positive effect on product evaluation. Due to the mixed design of this study, we included random intercepts for participants (SD = .94), random slopes for imbalance score (SD = 1.12) and rating variance (SD = .43) in our model. Given the importance of the sampling of stimuli (Judd, Westfall, and Kenny 2012), we also included random intercepts for the rating distributions that were used (SD = .32).

As proposed, displaying ratings in proportional bar graphs led to lower product evaluations (b = –.41, SE = .06, $p < .001$, 95% CI = [–.55, –.29]). We also found that the peak value had a positive impact on product evaluation (b = 1.38, SE = .20, $p < .001$, 95% CI = [1.00, 1.76]). As predicted in H_2, with an increase in the distribution's peak value, the negative impact of the graphical format on evaluation became weaker (b = .53, SE = .08, $p < .001$, 95% CI = [.38, .69]). Moreover, the effect of the graphical format was not moderated by the product type ($p = .932$).

We operationalized imbalance score by calculating the difference between the proportion of positive (i.e., sum of 4- and 5-star) and negative (i.e., sum of 1- and 2-star) ratings. Our analysis supported the positive impact of imbalance score on product evaluation (b = 1.63, SE = .21, $p < .001$, 95% CI = [1.23, 2.04]). More importantly, as predicted in H_3, the positive effect of imbalance score was weaker when ratings were displayed in the proportional format (b = –.19, SE = .05, $p < .001$, 95% CI = [–.31, –.10]), suggesting that consumers' perception of imbalance score is lower when a proportional bar graph is viewed. Consistent with prior research, we also found that average rating had a positive impact on evaluation (b = 1.63, SE = .12, $p < .001$, 95% CI = [1.38, 1.87]. While we expected a negative main effect of variance and a positive interaction effect with average rating, neither was significant (all *ps* > .400), possibly because the peak value and imbalance score included in our models had captured the effect of variance. See Table 2 and Appendix F for more details.

[INSERT TABLE 2 ABOUT HERE]

This study provided support for H_1, H_2, and H_3 by demonstrating that the negative effect of displaying ratings in the proportional format is generalizable to an extensive set of distributions and that this effect is moderated by the peak value. We also found that the positive impact of imbalance score on product evaluation is moderated by the graphical format in which rating distributions are displayed. We further showed that this effect holds for both small and relatively more significant purchases. Finally, while average rating had a positive impact on product evaluations, neither the variance nor its interaction with average rating predicted product evaluation.

Study 3: Ruling Out Alternative Explanations

We previously argued that the graphical presence of part-to-whole relationships is the underlying mechanism for the observed effect. However, in addition to visualizing rating information in a proportional or simple format, the bar graphs used in Study 2 (i.e., Figure 1's top panels) differed in other aspects that may explain our findings. In many cases, the lengths of the highlighted bars in simple bar graphs were greater than those of highlighted bars in proportional bar graphs. One possibility is that this difference in absolute sizes can be why participants evaluate products less positively when rating distributions are displayed in proportional bar graphs. However, this explanation is unlikely due to our between-subjects design, where participants saw only one graphical format and did not have a chance to compare bar lengths across the two graphical formats.

A second plausible explanation could be the differences in the lengths of the bars relative to the chart area. In other words, relative to the chart area, bars in a simple bar graph appear taller than do those in proportional bar graphs, and it could be argued that these differences in relative lengths had led to the observed effect in Study 2. A third possible explanation could be a difference in the number of graphical cues of the graphs and their appearances. Specifically, the presence of lightly shaded rectangles in proportional bar graphs used in Study 2 might have led to lower evaluations. While Study 3 tests these possibilities, it should be noted that differences in bar lengths and the number of graphical cues do not offer any predictions regarding our other propositions described above (i.e., H_2 and H_3). In this study, we also measured response time to test whether consumers need more time to process either of the graphical formats.

Method. We recruited 607 participants (69% female, mean age = 35.10 years) using Prolific (an online platform). The stimuli consisted of 21 distributions with an average rating of 3.0, 3.3, 3.6, 3.9, 4.2, 4.5, or 4.8 that were randomly selected from a pool of 130,000 randomly generated distributions (see more details in Appendix G). Participants were randomly assigned to one of the six format conditions displayed in Figure 3, where each person was shown rating distributions for seven TV sets in a randomized order.

[INSERT FIGURE 3 ABOUT HERE]

Including Formats II and V enabled us to test the impact of the presence of lightly shaded rectangles. The lightly shaded rectangles in Format II increase the number of graphical cues compared to Format I's blank spaces. Format V is a proportional bar graph because the placement of the bars' values adds a fixed reference point (i.e., 100% of the ratings) equivalent to the reference point offered in Format IV. Prior research has shown that consumers incorporate such reference points into their decision-making process (Thomas and Kyung 2019). Format III enabled us to test the impact of bars' visual saliency because it displays information in a simple format, while the bars appear the same size as those in the proportional formats (i.e., Formats IV and V). It should be noted that in Format III, the placement of bar values next to the bars hinders proportional judgments, but the salience of the chart border makes relative length judgments possible. Therefore, should our proposed mechanism be at play, the effect of using proportional (vs. simple) bar graphs on evaluation should be significant even after controlling for the presence or lack of gray rectangles and differences in bar lengths. Consistently, participants in Formats IV and V should evaluate the products less positively than those in Formats I, II, and III. The table condition (Format VI) was included as the control condition.

Results. We conducted a one-way repeated-measures ANOVA to test whether product evaluation was significantly different among the six conditions. As in previous studies, the format in which rating distributions were displayed had a significant effect on product evaluation ($F(5, 601) = 4.33$, $p = .001$, $\eta^2 = .03$). On average, product evaluations were lower at 5.61 (SD = 2.58) in the proportional bar graph (Formats IV and V combined) conditions compared to 6.00 (SD = 2.53) in the simple bar graph (Formats I, II, and III combined) conditions ($t(3547) = 4.48$, $p < .001$, Cohen's d = .15). Product evaluation in the table condition was 5.59 (SD = 2.41). Figure 4 displays product evaluations across the six conditions.

[INSERT FIGURE 4 ABOUT HERE]

A priori, we expected participants in both simple and proportional bar graph conditions to evaluate target items more favorably than those in the table condition. Research has shown that providing information in a graphical format can facilitate cognitive processes (Lurie and Mason 2007; Scaife and Rogers 1996), leading to higher evaluations. However, while participants in the simple bar graph conditions evaluated target products more positively than those in the table condition, there was no significant difference in product evaluation between the proportional bar graph and table conditions (p = .870). The fact that displaying ratings in a proportional format did not boost product evaluations suggests that offering graphical proportions cancels out the positive effect of using visual displays, supporting our central hypothesis (H_1).

We conducted four linear mixed-effects regression models to test the impact of bar lengths, the number of graphical cues (i.e., presence of lightly shaded rectangles), and the graphical presence of part-to-whole relationships on evaluations. As shown in Table 3, we included the graphical format (–1 = simple, +1 = proportional), its interaction with the peak value, and the imbalance score in all models except the base model. Similar to Study 2, we also included random intercepts for each participant (SD = .77), rating distribution (SD = .34), and random slopes for imbalance score (SD = .75) and rating variance (SD = .44) in our models. As expected, our proposed model performed better than the base, covariate-only, and full models according to the Akaike information criterion (AIC). This finding indicated that the proportional (vs. simple) aspect of graphical formats could better explain the observed differences in product evaluation across the conditions.

Based on this model, offering rating proportions graphically leads to a lower product evaluation (b = –.33, SE = .10, p = .001, 95% CI = [–.55, –.14]), supporting H_1. The significant and positive main effect of peak value on evaluation (b = 3.59, SE = .85, p < .001, 95% CI = [2.11, 5.11]) supported its impact on evaluation even when no graphical format was used. The negative and significant interaction between the display format and peak value (b = –1.58, SE = .43, p < .001, 95% CI = [–2.31, –.79]) indicated that the positive effect of peak value was smaller when distributions were displayed graphically compared with the tabular format. The nonsignificant main effect of imbalance score (b = .50, SE = .75, p = .509, 95% CI = [–.97, 2.16] showed that participants did not commit binary bias in the absence of graphical displays (i.e., in the tabular format). Its positive interaction with the display format (b = 1.13, SE = .24, p < .001, 95% CI = [.67, 1.55]), however, indicated that a higher imbalance score led to higher evaluations when ratings were displayed in a graph.

Importantly, this study also provided support for H_2 and H_3. As predicted in H_2, with an increase in the distribution's peak value, the negative impact of displaying ratings in proportional bar graphs on evaluation became weaker (b = .41, SE = .17, p = .015, 95% CI = [.10, .74]). Moreover, the positive effect of imbalance score when ratings were presented graphically was weaker in the proportional (vs. simple) format (b = –.19, SE = .09, p = .048, 95% CI = [–.35, .01]), suggesting that consumers were affected less by binary bias when a proportional bar graph was used.

Again, consistent with previous research, average rating had a positive effect on product evaluation (b = 1.96, SE = .45, p < .001, 95% CI = [.94, 2.86]). Similar to Study 2, rating variance and its interaction with average rating did not predict product evaluation *(all ps > .600).* Another important finding of this study was that consumers committed binary bias only when ratings were presented in a graphical format and not in a tabular format. This finding indicates that while using graphical displays in the context of online reviews may benefit firms and consumers, it also has drawbacks.

[INSERT TABLE 3 ABOUT HERE]

Processing fluency and response time. If consumers have more difficulty processing one type of information, they should require more time to process that piece of information (Westerman et al. 2015). In this experiment, we also measured the response time for each evaluation task. We conducted a one-way repeated-measures ANOVA to test the effect of the information presentation format on processing time. Our analysis revealed that the processing time was not significantly different across the six conditions ($F(5, 601) = .67$, $p = .646$). The response time in seconds for each task was 19.91 (SD = 23.37) in Format I, 17.79 (SD = 17.87) in Format II, 19.79 (SD = 39.1) in Format III, 19.16 (SD = 31.07) in Format IV, 17.05 (SD = 13.34) in Format V, and 17.21 (SD = 17.49) in Format VI. We further measured the correlation between the response time and product evaluation and did not find a significant correlation between them ($r = .005$, $p = .761$), undermining the role of processing fluency in the current context.

In sum, this study replicated the findings of Study 2 and showed that other visual differences and bar size could not explain the impact of graphical formats on consumers' evaluations. We also found no support for the processing fluency explanation. In the following study, we test the effect of graphical format on the perception of online consumer ratings.

Study 4: Consumers' Perception of Ratings in Proportional and Simple Bar Graphs

Thus far, we have provided converging evidence that consumers tend to evaluate products less positively when rating distributions are displayed in proportional bar graphs. We propose that this is because consumers tend to underestimate the value of large proportions (i.e., positive ratings) when ratings are displayed in proportional bar graphs. Thus, this study aimed to test the effects of the graphical format on the perception of ratings more rigorously.

Method. Participants (N = 297, 63% female, mean age = 35.17 years) were recruited via Prolific. In the instruction task, they were informed that they would be estimating the number of consumers who had given the product a specific rating score. To ensure that participants understood the task, in the instruction phase, they were asked to enter the number of 3-star reviews along with the total number of reviews displayed on a distribution graph. We then implemented a 2 (graphical format: simple vs. proportional) × 2 (valence: positive vs. negative ratings) × 4 (target value: the number of reviews with a specific rating score: 35 vs. 49 vs. 872 vs. 938) mixed design. The first two factors varied between subjects and the latter varied within subjects. Therefore, we randomly assigned participants to one of the four conditions where they estimated the number of reviews with a specific rating score. Distributions and their target rating scores are displayed in Figure 5.

As displayed in Figure 5, each participant saw four rating distributions—one at a time—with their total number of ratings but no information on the frequency or proportion of each star rating. They were then asked to provide their best and quick estimates of the number of ratings associated with the target bars, which represented 35, 49, 872, and 938 reviews. These corresponded to two small (3% and 4%) and two large (75% and 79%) proportions from distributions with a different number of total reviews (1,165, 1,230, 1,163, and 1,187, respectively). We expected consumers in the proportional bar graph condition to underestimate the number of ratings when the true number of ratings were high at 872 and 938 reviews. Moreover, for small values, 35 and 49, we expected participants in the proportional bar graph conditions to provide larger estimates.

[INSERT FIGURE 5 ABOUT HERE]

Results. Fourteen participants failed to complete the instruction task, leaving 283 participants. We divided the participants' estimates by the total number of reviews for each distribution to calculate the estimated proportions. We first conducted a three-way repeated-measures ANOVA to test the effect of the graphical format, true proportion, valence, and their interactions on proportion estimates. We applied the Greenhouse–Geisser corrections (Greenhouse and Geisser 1959) for lack of sphericity, and found that both the graphical format ($F(1, 279) = 13.03$, $p < .001$, $\eta^2_{partial} = .05$) and the true proportion ($F(1.83, 511.12) = 6687.44$, $p < .001$, $\eta^2_{partial} = .96$) had significant effects on participants' estimates. As expected, our analyses also showed that there was a significant interaction between the graphical format and true proportion ($F(1.83, 511.12) = 7.91$, $p = .001$, $\eta^2_{partial} = .03$). Moreover, there was no significant main effect of valence ($F(1, 279) = .15$, $p = .700$), or any other two- or three-way interactions among the three factors (all $ps > .250$).

Estimation of large proportions. When target values were large proportions (i.e., 75% and 79%), participants in the proportional bar graph conditions provided smaller estimates than those in the simple bar graph conditions. When the true proportion was 75% [79%], participants' mean estimates were 70.82% [72.47%] in the proportional bar graph conditions and were 76.08% [76.92%] in the simple bar graph conditions. The results of two t-test analyses showed that the estimates of these proportions were significantly lower in the proportional (vs. simple) bar graph conditions (true proportion 75%: $t(281) = –3.67$, $p < .001$; true proportion 79%: $t(281) = –3.29$, $p = .001$). This finding supports the idea that in proportional bar graphs, large proportions are underestimated.

Estimation of small proportions. When the true proportion was 3% [4%], participants' mean estimates were 5.15% [6.11%] (proportional bar graph) and 5.21% [6.41%] (simple bar graph), thereby overestimating small proportions. In other words, participants in both simple and proportional bar graph conditions overestimated smaller values. However, as expected, these overestimations were very small in both conditions (2% ~ 3%). Two t-test analyses found that there was no difference between the two conditions (true proportion 3%: $p = .930$; true proportion 4%: $p = .636$). These overestimations in both simple and proportional bar graph conditions support our proposed explanation that the lower evaluation of proportional bar graphs is due to the lower subjective value of large proportions.

Estimation errors. To test our explanation more rigorously, we also conducted a mixed-effects regression model to test the effect of the graphical format, true proportion, and their interaction effect on estimation errors. We calculated estimation errors by subtracting the true proportions from the estimated proportions and regressed it on the graphical format, true proportion, and their interaction effect. We also controlled for valence (i.e., whether the target value corresponded to positive or negative ratings) and response time (time spent estimating in seconds). We included random intercepts for each participant (SD = 1.59) and each stimulus (SD = .95) and random slopes for the true proportion (SD = 12.97; Table 4). A positive and statistically significant intercept of 2.52% (SE = .75, $p = .007$, 95% CI = [1.02, 3.91]) suggested that participants in all conditions overestimated smaller proportions. The negative main effect of the true proportion (b = –7.09, SE = 1.28, $p < .001$, 95% CI = [–9.59, –4.32]) and its interaction with the graphical format (b = –3.16, SE = .89 $p < .001$, 95% CI = [–5.00, –1.34]) suggested smaller estimation errors (i.e., smaller overestimations) for larger proportions in the simple bar graph conditions and an underestimation of larger proportions in the proportional bar graph conditions.

[INSERT TABLE 4 ABOUT HERE]

Response time. Similar to Study 3, we measured response time to test whether participants in one condition provided their estimates faster than others. Participants in the proportional and simple bar graph conditions spent 29.66 (SD = 32.89) and 27.31 (SD = 25.25) to estimate each target value, respectively. We conducted a similar three-way repeated-measures ANOVA and applied Greenhouse–Geisser corrections (Greenhouse and Geisser 1959) for lack of sphericity to test the effect of graphical format and its interaction with the true proportion and valence on estimation time. Neither was there a significant main effect of the graphical format ($F(1, 279) = .95$, $p = .330$), valence ($F(1, 279) = 1.15$, $p = .285$), or true proportion ($F(2.42, 674.64) = .51$, $p = .636$), nor were there any two- or three-way interactions among the three factors (all *ps* > .250).

In summary, this study demonstrated that consumers tend to overestimate small values and underestimate large values when a proportional bar graph is used. While the modest overestimation of small proportions was shared between the two formats, the underestimation of large proportions was unique to proportional bar graphs. These findings support the idea that the lower subjective value of large proportions (i.e., positive ratings) leads to a lower product evaluation when ratings are displayed in a proportional bar graph.

Study 5: The Effect of the Graphical Format on Choice: TripAdvisor Restaurant Page

In previous studies, we found that a distribution's graphical format significantly influences consumers' perception of ratings and their assessment of the products. The purpose of this study was to test whether the graphical format of rating distributions could also affect consumers' choices. As predicted in H_2, the negative impact of the graphical format should be strongest when a distribution has a small (vs. large) peak value. The distributions exhibited in Figure 6 have the same averages (4.4 out of 5). However, the peak value of the distribution for Keens Restaurant is smaller than that for Perkins Restaurant (50% vs. 78% 5-star ratings, respectively). Now, let us consider two choice sets, each consisting of two restaurants with rating distributions presented in simple or proportional bar graphs. Compared with the simple bar graph condition, the choice share of Keens Restaurant should drop in the proportional bar graph condition, as its rating distribution has a smaller peak value, and its evaluation is influenced more when its ratings are displayed in a proportional bar graph. In this experiment, we test this proposition in the presence of many other distracting pieces of information.

Method. We recruited 204 MTurk participants (39% female, mean age = 39.14 years). Participants were randomly assigned to the simple bar graph (Figure 6: choice set A) or proportional bar graph condition (Figure 6: choice set B) in a between-subjects design where participants were shown only the bottom panels of Figure 6. They were asked to imagine that they wanted to dine out and then selected the restaurant of their choice.

Results. Analysis revealed that out of 106 participants in the simple bar graph (choice set A) condition, 68 (64.2%) participants preferred Keens Restaurant compared with only 46 (46.9%) participants out of 98 in the proportional bar graph (choice set B) condition ($\chi^2(1) = 6.12$, $p = .013$). Logistic regression analysis using the graphical format as the only independent variable also showed that the choice share of this restaurant significantly decreased in the proportional bar graph condition (b = –.70, SE = .29, p = .014, 95% CI = [–1.27, –.15]). In summary, this study revealed that the graphical format of rating distributions significantly affects consumers' choices, even when rating distributions are only one of many available pieces of information.

[INSERT FIGURE 6 ABOUT HERE]

Summary of Additional Studies Reported in Appendices

In addition to the five studies reported here, we reported four more studies in the Appendix that tested other factors related to our propositions, including attention, the role of average rating, and the type of numerical information provided in the graphs. Appendix A was an eye-tracking study in which we measured pupil fixation to test the effect of graphical format on consumers' attention. In this study, we found that consumers relied more [less] on the graphical [textual] information when information was provided in a proportional (vs. simple) bar graph. We also found that the more time consumers spent on graphical cues in a proportional bar graph, the lower their evaluations were.

In Appendix B, in addition to testing the effect of graphical format on evaluation, we measured the perceived visual appeal of five information presentation formats. We observed the same pattern of results for product evaluation as in other studies but did not find any differences in the perceived visual appeal of the formats. This study provided further evidence against both relative and absolute length judgment explanations. In Appendix C, we provided additional evidence supporting the positive interaction between the graphical format and the peak value. Finally, Appendix D tested five graphical formats with different colors and different types of numerical information (i.e., the value of the bars displayed in frequencies or percentages), where the distribution graph took a small portion of the stimuli. This study validated previous findings and demonstrated that the negative effect of using a proportional bar graph persists, regardless of the type of numerical information and colors used to provide graphical part-to-whole relationships.

General Discussion

This work has focused on an important yet neglected aspect of online WOM. We systematically examined how displaying ratings in various bar graphs affects consumers' evaluations of products. Across nine studies (five reported in this article and four in the Appendices), we found strong evidence that consumers report a lower evaluation of an item when its rating distribution is displayed in a graphically proportional format. We attribute these findings to consumers' nature of processing involving the graphical displays of part-to-whole relationships (i.e., proportion judgments). Study 1 found that the same business is likely to receive different ratings on platforms that use different formats to display rating distributions. Utilizing an extensive set of rating distributions, Studies 2 and 3 established the negative effect of displaying ratings in proportional bar graphs.

However, the degree to which proportional bar graphs can lower consumers' evaluations is not equal across all distributions. In Studies 2 and 3, we found that the distribution's peak value moderates this effect, suggesting a smaller effect for distributions with greater peak values. We also found that other processes are influenced by the graphical format. The results of Studies 2 and 3 showed that consumers are less influenced by binary bias—considering different levels of star ratings as being equally positive (e.g., 4- and 5-star ratings) or equally negative (i.e., 1-

and 2-star ratings)—when ratings are displayed in proportional (vs. simple) bar graphs. Studies 3 and 4 did test our proposed mechanism and ruled out several crucial alternative explanations. Study 3 showed that neither the difference in relative and absolute sizes nor the presence of lightly shaded rectangles could have led to the results. In Study 4, we found that the negative effect of using a proportional bar graph is due mainly to the underestimation of large proportions. Study 5 established that the distributions' graphical format also influences consumer choice.

Implications. From a *theoretical perspective*, this research makes several contributions. First, we contribute to the large body of research on online WOM. The extant research on consumers' interpretations of online consumer ratings has been focused mainly on the impact of attributes such as average ratings, rating variance, and rating volume. Even the limited research that has studied the consequences of displaying rating distributions in a graphical format (e.g., Fisher, Newman, and Dhar 2018) did not differentiate between simple and proportional bar graphs. Our findings suggest that the format used to display the distribution of ratings influences consumers' perception of such ratings, their evaluation of products, and their choices. As a result, by identifying the graphical format of rating distributions as a key factor in consumers' evaluation, these findings enhance our understanding of how online ratings affect consumers.

Second, this research suggests that the findings of previous studies regarding the impact of displaying rating distributions may be moderated by the format in which distributions are presented. For example, we found that binary bias is less influential when proportional (vs. simple) bar graphs are used. Previous research has suggested that consumers prefer products with multimodal or unimodal rating distributions, depending on the context (Rozenkrants et al. 2017). Our findings in Study 5 suggest that consumers' preference for products with multimodal and unimodal distributions is also influenced by the graphical formats in which rating distributions are displayed.

Third, our work also contributes to graphical perception and visual information processing research (Cleveland and McGill 1984; Hegarty 2011; Lurie and Mason 2007; Talbot, Setlur, and Anand 2014). We augment this research stream by scrutinizing differences between interpretations of graphical formats in an online shopping context and by demonstrating how consumers make sense of different graphical representations of the same information. Our results are consistent with the idea that even small changes in relatively subtle aspects of graphical displays can affect task performance dramatically (e.g., Fischer 2000; Hegarty 2011; Lurie and Mason 2007). This research extends the understanding of how different perceptual tasks in decoding simple and proportional bar graphs can affect consumers' perception of ratings. Contributing to the extant research on proportion and length judgments, we also found that individuals tend to underestimate large proportions in the context of online consumer ratings.

From a *managerial perspective*, given the astronomical number of reviews being posted, the present work has important implications. Consumers are often overwhelmed in extracting meaningful insights from all the reviews submitted for an item. Summary information, such as rating distributions and average ratings, empowers consumers to evaluate and compare alternatives more efficiently. This work demonstrates that the relatively subtle features of online WOM, such as the graphical format of rating distributions, can substantially influence consumer judgments. Marketers should not ignore this aspect of online WOM, as it is one of the few attributes that they fully control and can modify at no additional cost. Our findings suggest that marketers can improve consumer evaluations of their offerings by displaying ratings in simple bar graphs, which is especially important for retailers or platforms that compete with other platforms. For instance, platforms such as Google Maps and TripAdvisor operate in the same industry and both derive advertising and transactional income by enabling business owners to advertise their services and allowing consumers to place online orders or make reservations. Our results suggest that platforms that display ratings in a proportional format (e.g., TripAdvisor) will be better off if they switch to a simple bar graph format.

To gain additional insights, we conducted a simulation. As TripAdvisor and Google use different graphical formats, we used the results of Study 2 to predict users' evaluations of the restaurants operating in New York City. Our simulation revealed that while TripAdvisor users' evaluations of restaurants in our dataset would be 6.51 (SD = .81) on average, Google users would evaluate the same restaurants 14% higher (M = 7.42, SD = .48). A paired t-test revealed that this difference was significant ($t(342) = 30.24$, $p < .001$). According to the National Restaurant Association, the industry generated $864.3 billion in sales in 2021 in the United States (FinanceOnline 2022). With this market size, even a small difference in consumers' evaluation of restaurants can have a substantial financial impact on both restaurants and review platforms.

It should be noted that according to our results, it is not recommended that all platforms display ratings in simple bar graphs. As suggested by previous research (Fisher et al. 2018) and supported by current findings, one consequence of displaying rating distributions is binary bias, where consumers group different levels of star ratings as either equally positive or equally negative. Now, let us consider a scenario in which committing binary bias has serious consequences or in which the difference between the utility of each star rating is substantial, as would be the case, say, when a 4-point versus 10-point rating scale is used. Alternatively, consider a product category for which a suboptimal choice can have substantial consequences.

For example, consumers and physicians may read patients' reviews for prescription drugs and their side effects on various platforms, such as AskAPatient.com. The cost of committing binary bias when deciding whether to take or prescribe a medication, either by a physician or a patient, would be greater than in the case of choosing a restaurant or a product such as a toaster. Thus, while the use of simple bar graphs may lead to higher product evaluations and potentially higher sales, they should not be used for product categories where a suboptimal choice may lead to potentially grave consequences. In such cases, platforms may even consider presenting rating distributions in tabular formats instead of any type of bar graph because—as Study 3 showed—consumers do not commit binary bias when tables are used. Indeed, the possibility of using graphical displays in an unethical manner to mislead consumers exists. Therefore, firms should use graphical displays cautiously and consider the ethical repercussions of their usage.

Limitations and future work. Although our findings are robust, they need to be viewed in light of two caveats. First, while we generated a population of 130,000 distributions and used samples from these, it was impossible to exhaust all possible distributions. Second, while consumers incorporate distribution information into their judgments, their evaluations are also influenced by other factors, such as prices, brand perception, and the textual content of reviews. For example, research has shown that the textual content of individual reviews can affect consumer decisions beyond other attributes of online WOM (Chen and Lurie 2013; Grewal and Stephen 2019; Kronrod and Danziger 2013; Rocklage and Fazio 2020), yet we did not investigate how the impact of using different graphical formats is influenced by the textual content of reviews. Moreover, consumers may put more (or less) weight on the rating distribution summary depending on its graphical format and the levels of other attributes (such as average rating, rating volume, price, or product category), possibly leading to different outcomes.

While we found consistent evidence supporting our proposition that graphical format influences consumer judgment and choice, we do not claim that all individuals are influenced by graphical formats in the same way. There can be individual variations, such as differences in numeracy levels or information processing styles (verbal vs. visual; Bagozzi 2008; Childers, Houston, and Heckler 1985), which can lead to different outcomes. Future work can also investigate additional graphical aspects of online WOM, such as the color and orientation of the distribution graphs. Whether consumers perceive vertical and horizontal bar graphs differently is also a worthwhile research question. Future research may similarly investigate the extent to which a bar graph's aspect ratio can influence consumers' perception in an online shopping context.

In conclusion, this research sheds new light on how consumer judgment is influenced by online rating distributions. We investigated the role of graphical displays in the context of online WOM and found that consumers' perceptions of ratings are affected by displays' graphical format, leading to diverse outcomes. We contribute to the marketing literature and graphical perception and information visualization research by showing how consumers perceive rating distributions differently depending on their graphical formats. In addition, this research has significant implications for consumers, academics and marketing practitioners by enhancing the understanding of rating distributions and their differential impact on consumer evaluations.

Appendices

Appendix A. Proportionalities and Information Diagnosticity: Measuring the Difference in Attention Through Eye Tracking

The format in which information is provided influences how consumers utilize various cues. As described in the main paper, research has shown that consumers tend to rely more on visual information when provided simultaneously with textual information, as they are perceived to be more diagnostic (Hegarty 2011; Lurie and Mason 2007). In the current context, therefore, prior research suggests that regardless of the graphical format used, consumers should engage in processing displayed bars (graphical cues) more than they do in textual information. The configuration of graphical cues is the only difference between the two formats, as the same information is provided. Therefore, any difference in consumers' judgments should be attributed to how consumers utilize various cues differently depending on the graphical format.

In the absence of textual information (i.e., value labels), consumers should engage in summation operations to estimate the value (frequency or proportion) of each star rating (Hollands and Spence 1998). On the other hand, in a proportional bar graph, the value of each star rating is displayed proportionally, and no additional mental calculations (summations and divisions) are needed. Since other than proportion judgments, no additional mental calculations are required to process graphical cues offered in proportional bar graphs, consumers' motivations for utilizing explicitly written values (i.e., value labels) should be lower when a proportional (vs. simple) bar graph is used. Therefore, we propose the following:

"*Consumers will rely more [less] on the graphical cues [explicitly written values] when distributions are displayed in the proportional (vs. simple) format."*

As we proposed earlier, proportion judgments are elicited through graphical cues. Paying more attention to graphical cues at the expense of explicit values should therefore increase biases in proportion judgments. In other words, spending more time processing graphical cues in a proportional bar graph does not allow consumers to adjust their biased perception of the ratings and therefore should make the effects of proportion judgments stronger. Accordingly, we further propose the following:

> "*The more consumers rely on graphical cues in a proportional bar graph, the lower will be their evaluation of the target item, while this should not be the case for simple bar graphs.*"

The purpose of Appendix A was to test the aforementioned predictions by demonstrating how consumers utilize graphical and textual information depending on the graphical format and how it may influence their evaluation of target items. Pupil fixation, which refers to brief instances in which the eyes focus on a given stimulus, is associated with the information being processed by the brain (Cian, Krishna, and Elder 2015) and is a reliable measure of the cognitive processing of information on which consumers act (Barbot et al. 2017). In this study, we captured participants' attention to test whether they rely more [less] on the graphical cues [bars' value labels] when processing a graph in a proportional (vs. simple) format and how their judgments are influenced by those differences.

Method. Undergraduate students (N = 159; 44% female; average age = 22.68 years) from a large U.S. university completed the study for partial course credit. We collected the data using a high frequency (120 Hz) eye-tracker (Tobii T120) that collected raw eye-movement data points every 8.3 milliseconds. The design was a single-factor—distribution graph format (Formats I vs. II vs. III, as shown in Figure 7)—between-subjects design, in which participants were asked to evaluate one product (a set of headphones). Formats I and III are simple and proportional bar graphs, respectively. Format II, on the other hand, is a variation of a simple bar graph in which blank spaces are replaced with lightly shaded rectangles. This graphical format does not provide part-to-whole relationships, but it does make the existing blank spaces more salient. Including this graphical format enabled us to test whether merely displaying gray rectangles would influence information processing and, as a result, pupil fixations. We asked one question to measure participants' willingness to buy headphones: "How likely would you be to buy this set of headphones?" (1 = "very unlikely," 10 = "very likely").

[INSERT FIGURE 7 ABOUT HERE]

Results. A one-way analysis of variance (ANOVA) examining the differences in participants' willingness to buy among the three graphical formats found significant differences ($F(2, 156) = 4.76$, $p = .038$). As expected, participants' willingness to buy the headphones was not different in the simple graphical (i.e., Formats I and II) conditions ($M_{\text{format I}} = 7.08$, $M_{\text{format II}} = 7.22$; $t(100) = .41$, $p = .681$). Therefore, we collapsed the data across these two conditions. In a planned contrast, participants in the proportional graph condition (Format III) reported lower willingness to buy the headphones than did those in the simple graphical (Formats I and II combined) conditions ($M_{\text{proportional bar graph}} = 6.40$ vs. $M_{\text{simple bar graph}} = 7.16$; $t(157) = –2.57$, $p = .011$, Cohen's d = .43), supporting H_1.

Processing Time. One explanation for the observed differences in the willingness to buy could be that processing a proportional bar graph requires more cognitive resources. If true, this could lead to lower processing fluency and lower product evaluations. Should this explanation be true, we should observe a significant difference in processing time across conditions (Westerman et al. 2015). In contrast, processing time was not found to be significantly different among the three conditions ($F(2, 156) = .81$, $p = .446$). However, as noted earlier, we expected participants in the proportional format (vs. simple format) to rely more [less] on the graphical cues [bar value labels] to assess the ratings. Consistent with this prediction, our analysis of gaze times indeed found that participants in the proportional format spent significantly more of their time looking at the graphical cues (highlighted bars and blank/shaded spaces combined) than did those in the simple bar graph conditions ($M_{\text{proportional bar graph}} = 66.45\%$, $M_{\text{simple bar graph}} = 57.57\%$; $t(157) = 3.30$, $p = .001$). This difference was driven mainly by the proportion of time participants in the proportional bar graph condition spent on lightly shaded rectangles, which elicited proportional judgments ($M_{\text{proportional bar graph}} = 37.80\%$, $M_{\text{simple bar graph}} = 18.96\%$; $t(157) = 8.26$, $p < .001$). On the other hand, participants in the proportional format spent significantly less of their time looking at the written values ($M_{\text{proportional bar graph}} = 13.88\%$) than did those in the simple bar graph conditions ($M_{\text{simple bar graph}} = 22.61\%$; $t(157) = -3.66$, $p < .001$). These results support our proposition that consumers place more weight on graphical cues at the expense of explicitly written values when processing a graph in a proportional format.

We argued earlier that the differences in the graphical cues between the proportional and simple format result in lower evaluations in the proportional format. Therefore, differences in evaluations should be greater as consumers place more weight on graphical cues, as their judgments are adjusted by explicitly written values to a lesser extent. To test this, we regressed mean-centered willingness to buy on the graphical format (proportional vs. simple), the mean-centered proportion of time gazed at graphical cues, and their interaction effect. The main effect of the graphical format was significant and negative (b = –.34, SE = .17, $p = .048$, CI = [–.672, –.003]), supporting our primary proposition that consumers evaluate a proportional bar graph less favorably than a simple bar graph. The main effect of time gazing at graphical cues was not significant (b = .07, SE = .10, p = .483, CI = [–.12, .26]), suggesting that the amount of time spent on graphical cues did not affect consumers' willingness to buy in simple bar graph conditions. As expected, we found a negative interaction effect between the time gazing at graphical cues and the graphical format such that higher gazing time at graphical cues in the proportional format led to a lower willingness to buy (b = –.35, SE = .17, p = .047, CI = [–.687, –.004]; see Figure 8).

[INSERT FIGURE 8 ABOUT HERE]

Appendix B. Role of Proportionalities and Relative Length Judgments

The primary aim of this study was to assess whether consumers judge the length of the bars relative to the chart border to evaluate a rating distribution rather than graphical proportions. This study also examined whether consumers perceive one form of information presentation as more visually appealing than the other.

Method. Three hundred seventy-seven participants (73% female, mean age = 26.87 years) were recruited using an online platform (Prolific). Participants were randomly assigned to one of the five distribution format conditions displayed in Figure 9A and asked to evaluate a brand-new coffee maker. We used the same items as we did in Studies 2 and 3 to measure product evaluation. After reporting their product evaluations, participants reported the visual appeal of the stimuli and answered basic demographic questions.

Stimuli. Formats I and IV in Figure 9A are original proportional and simple bar graphs used in previous studies. Format II is also a proportional bar graph where we displace value labels to make it more comparable to Format III. Compared to Format I, we expected participants who saw the rating distribution in Format II to report similar evaluations, as the presence of gray rectangles still triggers proportion judgments. The absolute and relative lengths of the bars are exactly the same in Formats I, II, and III. However, in Format III, we interfere with proportion judgments by simultaneously removing gray rectangles and displacing explicitly written values. If lower evaluations of proportional bar graphs are caused by bars' smaller relative or absolute lengths, participants who see the distribution in Format III should also report a lower evaluation than those who see the distribution in Format IV. However, if proportion judgments biases are at play, we should not observe any difference in evaluations between participants in Formats III and IV conditions, despite differences in bar lengths. Finally, we included the table condition as the control condition in which no graphical cues were provided.

Results. We conducted two one-way ANOVAs to test whether product evaluation and the perceived visual attractiveness were significantly different among the five conditions. As in previous studies, the graphical format of the rating distributions was found to have a significant effect on product evaluation ($F(4, 372) = 6.16$, $p < .001$). Visual appeal, however, was not perceived to be significantly different across the five conditions ($F(4, 372) = 1.42$, $p = .228$). Consistent with our theorizing, participants who saw the rating distribution in both Formats III and IV evaluated the coffee maker more favorably than did those seeing the distribution in Formats I and II ($M_{III} = 6.80$, $M_{IV} = 6.96$; $M_{I} = 6.10$, $M_{II} = 5.93$; all $ps <= .01$). The higher evaluation of ratings in Format III than in Formats I and II is theoretically important, as it supports the idea that proportion judgments dampen the consumer evaluations in the proportional format rather than the absolute or relative sizes of the bars. Finally, it should be noted that the difference in product evaluations between Formats III and IV and the control condition was not significant ($M_{table} = 6.92$; all $ps > .6$; see Figure 9).

[INSERT FIGURE 9 ABOUT HERE]

Appendix C. Average Rating, Peak Value, and Graphical Format Interaction Effects

Although we predicted that the peak value would moderate the negative effect of using proportional bar graphs and Studies 2 and 3 supported our prediction, we did not predict consumers' evaluations to be moderated by the average rating. Our post hoc analyses in the abovementioned studies also did not support such an interaction effect. However, we conducted this straightforward between-subjects study to test whether the effect of the graphical format was moderated by average rating or peak value.

Method. We recruited 721 participants (42% female, average age = 37.95 years) from MTurk in return for a small payment. This study involved a 2 (graphical format: simple vs. proportional bar graph) × 2 (average rating: 4.4 vs. 3.4 out of 5 stars) × 2 (peak value: high vs. low) between-subjects design, as shown in Figure 11. Participants were told that the distribution displayed ratings for a coffee maker, and using the scale used in Studies 2 and 3, each participant evaluated one product. As mentioned earlier, we expected this more straightforward design to shed light on how these factors influence participants' evaluations.

Results. We conducted a 3-way between-groups ANOVA (Table 5) to examine the main effects and interactions of the graphical format, average rating, and peak value on product evaluation. Consistent with our previous findings, product evaluation was found to be significantly influenced by the graphical format ($F(1, 713) = 4.33$, p =.038), with proportional graphs leading to a lower product evaluation (M = 7.34) than simple bar graphs (M = 7.61). The main effects of average rating ($F(1, 713) = 82.76$, $p < .001$) and peak value ($F(1, 713) = 8.53$, $p = .004$) were also significant. While the interaction between the graphical format and peak value was significant ($F(1, 713) = 3.68$) $p = .055$), that between the graphical format and average rating was not ($F(1, 713) = .17$, $p = .676$).

[INSERT FIGURE 10 ABOUT HERE]

[INSERT FIGURE 11 ABOUT HERE]

[INSERT TABLE 5 ABOUT HERE]

Appendix D. Variations in Numerical Information, Colors, and Additional Information

The objectives of this study were twofold: we wanted to test 1) whether the effect of the graphical format on evaluation holds when additional pieces of information compete for consumers' attention and 2) whether a match or mismatch between the graphical format and explicitly written values influences the results. In other words, we wanted to test how using written raw numbers (i.e., frequencies) versus percentages for the value of the bars and the color of empty spaces may affect product evaluation.

Method. We recruited 400 participants (44% female; average age = 39.31 years) using MTurk. The same pool of 21 distributions and measurement scale employed in Study 3 were used. We randomly assigned participants to one of the five formats exhibited in Figure 12 in a between-subjects design, resulting in 105 stimuli. We designed our stimuli to resemble a product page on Amazon.com. Formats I, II, III, and IV allowed us to test the impact of the type of numerical information provided in a distribution graph (written frequencies vs. percentages for the value of each bar). Format V was included to test the impact of the color of the additional cues in proportional bar graphs. Within subjects, each participant evaluated seven different products (rechargeable LED lamp, face masks, charcoal grill, coffee maker, toaster, air circulator, and microwave oven) with different rating distributions in a randomized order.

Results. We conducted a linear mixed-effects regression analysis. We used average rating, variance of ratings, their interaction, color, the distribution graph format (–1 = simple bar graphs, +1 = proportional bar graphs), type of written values (–1 = frequencies; +1 = percentages), and type of written values × graphical format as fixed effects. Random intercepts for participants, products, and distributions in addition to random slopes for imbalance score and variance were included (see Table 6). Participants evaluated the products less favorably when ratings distributions were presented in proportional formats (b = –.30, SE =.13, p =.021, 95% CI = [–.55, –.04]). The average product evaluation in the proportional bar graph conditions was 6.24 (SD = 1.97) compared with 6.68 (SD = 2.01) in the simple bar graph conditions (Cohen's d = .22). Using a different color (b = –.05, p = .642) or different types of value labels (b = .05, p = .531) did not influence participants' judgments significantly. More importantly, the influence of the graphical format was not moderated by the type of value labels (b = –.01, p = .899). This finding supports the idea that the negative effect of using a proportional bar graph holds even when (1) other information (e.g., product description, rating volume, and graphical and verbal presentation of the means) is provided and (2) bar values are provided in frequencies and percentages.

[INSERT FIGURE 12 ABOUT HERE]

[INSERT TABLE 6 ABOUT HERE]

[INSERT FIGURE 13 ABOUT HERE]

Appendix E. Attention Check Used at the Beginning of MTurk and Prolific Studies

Below are several questions related to bamboo. We want to test the level of general knowledge about bamboo among the general American population. Please read the questions and select one answer. When you select the answer, always choose the answer that says **'none of the**

above’ instead of clicking on the right answer. In other words, please ignore the right answer, and just click on the answer choice that says ‘none of the above’.

What is the growth rate of a normal healthy bamboo plant?

about 3-10 cm in 24 hours	about 15-20 cm in 24 hours	about 20-25 cm in 24 hours	none of the above
○	○	○	○

How long does bamboo live?

about 10-20 years	none of the above	about 20-40 years	about 40-60 years
○	○	○	○

What country eats food made with bamboo?

China	Nepal	none of the above	Both
○	○	○	○

Appendix F. Distributions used in Study 2 and Additional Tested Models

[INSERT TABLE 7 ABOUT HERE]

[INSERT TABLE 8 ABOUT HERE]

[INSERT TABLE 9 ABOUT HERE]

[INSERT TABLE 10 ABOUT HERE]

[INSERT TABLE 11 ABOUT HERE]

[INSERT TABLE 12 ABOUT HERE]

[INSERT TABLE 13 ABOUT HERE]

[INSERT TABLE 14 ABOUT HERE]

Appendix G: Distributions Used in Study 3 and Additional Tested Models

[INSERT TABLE 15 ABOUT HERE]

[INSERT TABLE 16 ABOUT HERE]

[INSERT TABLE 17 ABOUT HERE]

[INSERT TABLE 18 ABOUT HERE]

[INSERT TABLE 19 ABOUT HERE]

Tables

TABLE 1: STUDY 1 MIXED-EFFECTS REGRESSION RESULTS

Predictor	*Estimates*	*SE*	*p-value*
(Intercept)	4.39	.03	<.001
Platform (TripAdvisor)	–.18	.01	<.001
Michelin Star (yes)	.13	.03	<.001
Price	.001	.01	.970
Platform × Michelin Star	–.03	.01	<.001
Platform × Price	.01	.003	<.001
Random effects	Variable		SD
Grouping variable: Restaurant	Intercept		.19
AIC			2,664,172
BIC			2,664,266

TABLE 2: MIXED-EFFECTS REGRESSION RESULTS IN STUDY 2

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–2.34	.39	<.001	–3.16	–1.53
Average Rating	1.63	.12	<.001	1.38	1.87
Variance	–.07	.25	.779	–.56	.37
Average Rating × Variance	.05	.07	.429	–.07	.19
Imbalance Score	1.63	.21	<.001	1.23	2.04
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.41**	**.06**	**<.001**	**–.55**	**–.29**
Peak Value	1.38	.20	<.001	1.00	1.76
Graphical Format × Imbalance Score (H_3)	**–.19**	**.05**	**<.001**	**–.31**	**–.10**
Graphical Format × Peak Value (H_2)	**.53**	**.08**	**<.001**	**.38**	**.69**
Graphical Format × Product	.01	.07	.932	–.12	.15
Random effects				Variable	SD
Grouping variable: Subject				Intercept	.94
		Variance		Slope	.43
		Imbalance Score		Slope	1.12
Grouping variable: Stimulus				Intercept	.32
AIC					40,719
BIC					40,837

~ –1 = simple, +1 = proportional

TABLE 3: STATISTICAL SUMMARY OF THE MIXED-EFFECTS REGRESSION MODELS TESTED IN STUDY 3

Outcome variable: Product evaluation					
Predictor		Base model *b* (*SE*)	Covariate-only model *b* (*SE*)	Proposed model *b* (*SE*)	Full model *b* (*SE*)
(Intercept)		–4.65 (1.35)***	–4.65 (1.35)***	–4.97 (1.36)***	–4.94 (1.36)***
Average Rating		1.95 (.45)***	1.96 (.45)***	1.96 (.45)***	1.96 (.45)***
Variance		–.19 (.86)	–.18 (.86)	–.18 (.86)	–.18 (.86)
Average Rating × Variance		.11 (.22)	.11 (.22)	.11 (.22)	.11 (.22)
Imbalance Score		.69 (.75)	.69 (.75)	.50 (.75)	.50 (.75)
Peak Value		3.17 (.83)***	3.17 (.83)***	3.59 (.85)***	3.59 (.85)**
Display Format (graph vs. table)		.52 (.23)**	.53 (.24)**	.78 (.25)***	.81 (.26)***
Display Format × Imbalance		.98 (.23)***	.98 (.23)***	1.13 (.24)***	1.13 (.24)***
Display Format × Peak Value		–1.25 (.40)***	–1.25 (.40)***	–1.58 (.43)***	–1.58 (.43)***
Graphical Format ~ (H_1)		–	–	**–.33 (.10)*****	–.29 (.11)**
Graphical Format × Imbalance Score (H_3)		–	–	**–.19 (.09)****	–.19 (.09)**
Graphical Format × Peak Value (H_2)		–	–	**.41 (.17)****	.41 (.17)**
Bar Size		–	.25 (.10)**	**–**	.06 (.14)
Gray Rectangles		–	–.26 (10)**	**–**	–.18 (.11)*
Random effects	Variable	SD	SD	SD	SD
Grouping variable: Subject	Intercept	.78	.77	**.77**	.77
Variance	Slope	.45	.45	**.44**	.45
Imbalance Score	Slope	.76	.75	**.75**	.75
Grouping variable: Stimulus	Intercept	.34	.34	**.34**	.34
AIC		14,992	14,985	**14,980**	14,981
BIC		15,081	15,087	**15,088**	15,102

* $p < .1$, ** $p < .05$, *** $p < .01$
~ –1 = simple, +1 = proportional

TABLE 4: MIXED-EFFECTS REGRESSION RESULTS IN STUDY 4

Outcome variable: Estimation error					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	2.52	.75	.007	1.01	3.91
Graphical Format~	.01	.26	.961	–.53	.53
True Proportion	**–7.09**	**1.28**	**<.001**	**–9.59**	**–4.32**
Valence (positive vs. negative)	–.03	.92	.978	–1.90	1.80
Graphical Format × True Proportion	**–3.16**	**.89**	**<.001**	**–5.00**	**–1.34**
Estimation Time	.00	.89	.971	–.01	.01
Random effects			Variable	SD	
Grouping variable: Subject			Intercept	1.59	
		True Proportion	Slope	12.97	
Grouping variable: Stimulus			Intercept	.95	
AIC				7,754	
BIC				7,805	

TABLE 5: TYPE III ANOVA RESULTS IN APPENDIX C

Predictor	SS	DF	F-value	p-value
Graphical Format	**13.98**	**1**	**4.33**	**.038**
Average Rating	267.40	1	82.76	<.001
Peak Value	27.56	1	8.53	.004
Graphical Format × Average Rating	.56	1	.17	.676
Graphical Format × Peak Value	**11.90**	**1**	**3.68**	**.055**
Average Rating × Peak Value	10.33	1	3.20	.074
Graphical Format × Average Rating × Peak Value	**.04**	**1**	**.01**	**.912**
Residuals	2303.70	713		

TABLE 6: APPENDIX D FULL MODEL

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	1.88	.76	.023	.30	3.48
Average Rating	.98	.26	.001	.43	1.48
Variance	.22	.49	.655	–.65	1.31
Average Rating × Variance	–.02	.12	.854	–.30	.21
Imbalance Score	.66	.42	.132	–.13	1.58
Peak Value	.34	.43	.448	–.46	1.16
Color (red vs. gray)	–.05	.10	.642	–.25	.16
Value Labels	**.05**	**.07**	**.531**	**–.10**	**.21**
Graphical Format ~ (H_1)	–.29	.13	.021	–.55	–.04
Graphical Format × Peak Value (H2)	.35	.22	.110	–.04	.80
Graphical Format × Imbalance Score (H3)	–.19	.13	.124	–.44	.06
Graphical Format × Value Labels	**–.01**	**.07**	**.899**	**–.14**	**.13**

Random Effects		Predictor variable	SD
Grouping variable: Subject		Intercept	.93
	Variance	Slope	.49
	Imbalance Score	Slope	1.00
Grouping variable: Product		Intercept	.00
	Stimulus	Intercept	1.37
AIC			10725
BIC			10832

~ –1 = simple, +1 = proportional

TABLE 7: AVERAGE RATINGS AND VARIANCE OF DISTRIBUTIONS USED IN STUDY 2

Average Rating (±.02)	3.0	3.2	3.4	3.6	3.8	4.0	4.2	4.4	4.6	4.8
Variance	.04	.28	.50	.30	.48	.17	.22	.24	.34	.17
	.08	.35	.58	.38	.56	.19	.34	.44	.54	.22
	.11	.40	.70	.54	.69	.33	.34	.48	.65	.29
	.24	.46	.77	.65	.71	.58	1.01	.50	.67	.33
	.38	.75	.77	.69	.75	.64	1.02	.61	.71	.34
	.38	.77	.93	.79	.83	.78	1.10	.64	.75	.35
	.89	.97	1.17	.91	1.04	1.08	1.14	.77	.75	.36
	1.09	1.21	1.17	.97	1.13	1.22	1.34	.83	.75	.42
	1.25	1.21	1.17	.99	1.66	1.22	1.53	1.09	.81	.43
	1.28	1.33	1.19	1.01	1.70	1.24	1.55	1.11	.85	.44
	1.68	1.35	1.36	1.15	1.88	1.26	1.57	1.13	.89	.44
	1.70	1.40	1.88	1.19	1.99	1.30	1.58	1.23	.89	.45
	1.72	1.53	2.00	1.25	2.01	1.34	1.59	1.29	.93	.45
	1.91	1.76	2.00	1.33	2.07	1.67	1.65	1.43	.99	.46
	2.00	1.79	2.08	1.44	2.19	1.73	1.66	1.47	1.05	.47
	2.01	1.82	2.22	1.57	2.48	1.77	1.77	1.47	1.07	.53
	2.59	1.83	2.35	1.62	2.48	1.81	1.86	1.47	1.11	.55
	2.59	2.01	2.50	1.64	2.81	2.17	2.05	1.48	1.13	.57
	2.94	3.37	2.70	1.64	2.89	2.62	2.16	1.76	1.27	.61
	3.18	3.97	2.85	1.90	3.39	2.66	2.46	1.96	1.42	.71

TABLE 8: STUDY 2 REGRESSION MODEL W1

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–2.34	.39	<.001	–3.16	–1.52
Average Rating	1.63	.12	<.001	1.38	1.87
Variance	–.07	.25	.779	–.56	.37
Average Rating × Variance	.05	.07	.430	–.07	.19
Imbalance Score	1.63	.21	<.001	1.23	2.04
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.40**	**.14**	**.004**	**–.68**	**–.11**
Peak Value	1.38	.20	<.001	1.00	1.76
Graphical Format × Imbalance Score (H_3)	**–.19**	**.09**	**.030**	**–.38**	**–.01**
Graphical Format × Peak Value (H_2)	**.53**	**.08**	**<.001**	**.37**	**.69**
Graphical Format × Product	.01	.07	.932	–.12	.15
Graphical Format × Average Rating	.00	.04	.941	–.10	.08
Random effects			Variable	SD	
Grouping variable: Subject			Intercept	.94	
	Variance		Slope	.43	
	Imbalance Score		Slope	1.12	
Grouping variable: Stimulus			Intercept	.32	
AIC				40,721	
BIC				40,846	

TABLE 9: STUDY 2 REGRESSION MODEL W2

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–2.35	.39	<.001	–3.18	–1.54
Average Rating	1.63	.12	<.001	1.38	1.88
Variance	–.05	.25	.828	–.54	.38
Average Rating × Variance	.05	.07	.474	–.07	.18
Imbalance Score	1.62	.21	<.001	1.23	2.04
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.33**	**.14**	**.017**	**–.61**	**–.05**
Peak Value	1.39	.20	<.001	1.00	1.76
Graphical Format × Imbalance Score (H_3)	**–.15**	**.09**	**.094**	**–.34**	**.03**
Graphical Format × Peak Value (H_2)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Product	.00	.07	.951	–.12	.15
Graphical Format × Average Rating	.07	.04	.126	–.03	.15

Random effects		Variable	SD
Grouping variable: Subject		Intercept	.94
	Variance	Slope	.43
	Imbalance Score	Slope	1.11
Grouping variable: Stimulus		Intercept	.32
AIC			40,765
BIC			40,883

TABLE 10: STUDY 2 REGRESSION MODEL W3

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–1.86	.43	<.001	–2.75	–.95
Average Rating	1.78	.14	<.001	1.52	2.06
Variance	–.54	.27	.042	–1.07	–.02
Average Rating × Variance	.14	.07	.044	.01	.29
Imbalance Score	1.54	.23	<.001	1.09	2.00
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.34**	**.14**	**.016**	**–.62**	**–.05**
Peak Value	-	-	-	-	-
Graphical Format × Imbalance Score (H_3)	**–.15**	**.09**	**.093**	**–.34**	**.03**
Graphical Format × Peak Value (H_2)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Product	.00	.07	.957	–.12	.15
Graphical Format × Average Rating	.07	.04	.123	–.03	.15

Random effects		Variable	SD
Grouping variable: Subject		Intercept	.94
	Variance	Slope	.43
	Imbalance Score	Slope	1.11
Grouping variable: Stimulus		Intercept	.36
AIC			40,808
BIC			40,919

TABLE 11: STUDY 2 REGRESSION MODEL W4

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–1.86	.43	<.001	–2.76	–.95
Average Rating	1.78	.14	<.001	1.52	2.06
Variance	–.54	.27	.043	–1.07	–.02
Average Rating × Variance	.14	.07	.045	.01	.29
Imbalance Score	1.54	.23	<.001	1.09	2.00
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.17**	**.10**	**.085**	**–.37**	**.02**
Peak Value	-	-	-	-	-
Graphical Format × Imbalance Score (H_3)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Peak Value (H_2)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Product	.00	.07	.957	–.12	.15
Graphical Format × Average Rating	.01	.02	.816	–.04	.05
Random effects			Variable	SD	
Grouping variable: Subject			Intercept	.94	
	Variance		Slope	.43	
	Imbalance Score		Slope	1.11	
Grouping variable: Stimulus			Intercept	.36	
AIC				40,809	
BIC				40,912	

TABLE 12: STUDY 2 REGRESSION MODEL W5

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–2.36	.39	<.001	–3.18	–1.55
Average Rating	1.64	.12	<.001	1.39	1.88
Variance	–.52	.25	.835	–.54	.39
Average Rating × Variance	.05	.07	.480	–.07	.18
Imbalance Score	1.62	.21	<.001	1.22	2.04
Product (Stapler)	.29	.07	<.001	.15	.42
Graphical Format~ (H_1)	**–.17**	**.10**	**.086**	**–.36**	**.03**
Peak Value	1.39	.20	<.001	1.00	1.76
Graphical Format × Imbalance Score (H_3)	-	-	-	-	-
Graphical Format × Peak Value (H_2)	-	-	-	-	-
Graphical Format × Product	.00	.07	.951	–.12	.15
Graphical Format × Average Rating	.01	.02	.830	–.04	.05
Random effects			Variable	SD	
Grouping variable: Subject			Intercept	.94	
	Variance		Slope	.43	
	Imbalance Score		Slope	1.11	
Grouping variable: Stimulus			Intercept	.32	
AIC				40,767	
BIC				40,877	

TABLE 13: STUDY 2 REGRESSION MODEL W6

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–4.17	.36	<.001	–4.82	–3.46
Average Rating	2.40	.09	<.001	2.23	2.56
Variance	–.24	.28	.395	–.81	.29
Average Rating × Variance	.06	.07	.461	–09	.21
Imbalance Score	-	-	-	-	-
Product (Stapler)	.15	.06	.010	.03	.27
Graphical Format~ (H_1)	**–.14**	**.09**	**.109**	**–.30**	**.05**
Peak Value	1.29	.22	<.001	.85	1.71
Graphical Format × Imbalance Score (H_3)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Peak Value (H_2)	**.47**	**.08**	**<.001**	**.31**	**.63**
Graphical Format × Product	.01	.06	.837	–.12	.13
Graphical Format × Average Rating	–.08	.02	<.001	–.14	–.04
Random effects			Variable	SD	
Grouping variable: Subject			Intercept	.85	
	Variance		Slope	.45	
Grouping variable: Stimulus			Intercept	.37	
AIC				41,236	
BIC				41,340	

TABLE 14: STUDY 2 REGRESSION MODEL W7

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–3.63	.38	<.001	–4.30	–2.88
Average Rating	2.50	.09	<.001	2.32	2.68
Variance	–.67	.29	.022	–1.20	–.11
Average Rating × Variance	.14	.08	.074	–.01	.30
Imbalance Score	-	-	-	-	-
Product (Stapler)	.15	.06	.011	.03	.27
Graphical Format~ (H_1)	**–.14**	**.09**	**.120**	**–.30**	**.06**
Peak Value	-	-	-	-	-
Graphical Format × Imbalance Score (H_3)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Peak Value (H_2)	**-**	**-**	**-**	**-**	**-**
Graphical Format × Product	.01	.06	.877	–.12	.13
Graphical Format × Average Rating	.00	.02	.803	–.05	.03

Random effects		Variable	SD
Grouping variable: Subject		Intercept	.85
	Variance	Slope	.44
Grouping variable: Stimulus		Intercept	.41
AIC			41,295
BIC			41,384

TABLE 15: AVERAGE RATING AND VARIANCE OF DISTRIBUTIONS USED IN STUDY 3

Average Rating (±.02)	3.0	3.3	3.6	3.9	4.2	4.5	4.8
Variance	.93	1.17	.30	.54	.83	.52	.33
	1.72	1.23	.91	1.58	1.50	.62	.44
	2.38	1.96	1.39	2.09	2.25	1.42	.70

TABLE 16: STUDY 3 BASE MODEL

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–4.65	1.35	.002	–7.18	–1.85
Average Rating	1.95	.45	<.001	.94	2.86
Variance	–.19	.86	.831	–1.82	1.48
Average Rating × Variance	.11	.22	.610	–.32	.54
Imbalance Score	.69	.75	.366	–.72	2.38
Peak Value	3.17	.83	<.001	1.65	4.68
Display Format (graph vs. table)	.52	.23	.025	.07	.92
Display Format × Imbalance	.98	.23	<.001	.56	1.38
Display Format × Peak Value	–1.25	.40	.002	–1.96	–.47
Random Effects			Predictor variable	SD	
Grouping variable: Subject			Intercept	.78	
		Variance	Slope	.45	
		Imbalance Score	Slope	.76	
Grouping variable: Stimulus			Intercept	.34	
AIC				14,992	
BIC				15,081	

TABLE 17: STUDY 3 COVARIATE-ONLY MODEL

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–4.65	1.35	.002	–7.18	–1.85
Average Rating	1.96	.45	<.001	.94	2.86
Variance	–.18	.86	.832	–1.82	1.49
Average Rating × Variance	.11	.22	.610	–.32	.54
Imbalance Score	.69	.75	.366	–.72	2.38
Peak Value	3.17	.83	.001	1.65	4.69
Display Format (graph vs. table)	.53	.24	.028	.04	.95
Display Format × Imbalance	.98	.23	<.001	.56	1.38
Display Format × Peak Value	–1.25	.40	.002	–1.96	–.47
Gray Rectangles	–.26	.10	.011	–.44	–.05
Bar Size	.25	.10	.013	.05	.44
Random Effects			Predictor variable	SD	
Grouping variable: Subject			Intercept	.77	
		Variance	Slope	.45	
		Imbalance Score	Slope	.75	
Grouping variable: Stimulus			Intercept	.34	
AIC				14,985	
BIC				15,087	

TABLE 18: STUDY 3 PREDICTED MODEL

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–4.97	1.36	.001	–7.52	–2.18
Average Rating	1.96	.45	<.001	.94	2.86
Variance	–.18	.86	.833	–1.82	1.49
Average Rating × Variance	.11	.22	.611	–.32	.54
Imbalance Score	.50	.75	.509	–.97	2.16
Peak Value	3.59	.85	<.001	2.11	5.11
Display Format (graph vs. table)	.78	.25	.001	.32	1.21
Display Format × Imbalance	1.13	.24	<.001	.67	1.55
Display Format × Peak Value	–1.58	.43	<.001	–2.31	–.79
Graphical Format ~ (H_1)	–.33	.10	.001	–.55	–.14
Graphical Format × Peak Value (H2)	.41	.17	.015	.10	.74
Graphical Format × Imbalance Score (H3)	–.19	.09	.048	–.35	.01

Random Effects		Predictor variable	SD
Grouping variable: Subject		Intercept	.77
	Variance	Slope	.44
	Imbalance Score	Slope	.75
Grouping variable: Stimulus		Intercept	.34
AIC			14,980
BIC			15,088

TABLE 19: STUDY 3 FULL MODEL

Outcome variable: Product evaluation					
Predictor	*Estimates*	*SE*	*p-value*	*95% CI*	
(Intercept)	–4.94	1.36	.001	–7.44	–2.13
Average Rating	1.96	.45	<.001	.94	2.86
Variance	–.18	.86	.833	–1.81	1.49
Average Rating × Variance	.11	.22	.611	–.32	.54
Imbalance Score	.50	.75	.510	–.97	2.16
Peak Value	3.59	.85	<.001	2.11	5.11
Display Format (graph vs. table)	.81	.26	.002	.31	1.28
Display Format × Imbalance	1.13	.24	<.001	.67	1.55
Display Format × Peak Value	–1.58	.43	<.001	–2.31	–.79
Gray Rectangles	–.18	.11	.098	–.39	.03
Bar Size	.06	.14	.685	–.24	.35
Graphical Format ~ (H_1)	–.29	.11	.007	–.53	–.08
Graphical Format × Peak Value (H2)	.41	.17	.015	.10	.74
Graphical Format × Imbalance Score (H3)	–.19	.09	.048	–.35	.01

Random Effects		Predictor variable	SD
Grouping variable: Subject		Intercept	.77
	Variance	Slope	.45
	Imbalance Score	Slope	.75
Grouping variable: Stimulus		Intercept	.34
AIC			14,981
BIC			15,102

Figures

FIGURE 1: PROPORTIONAL AND SIMPLE BAR GRAPH EXAMPLES

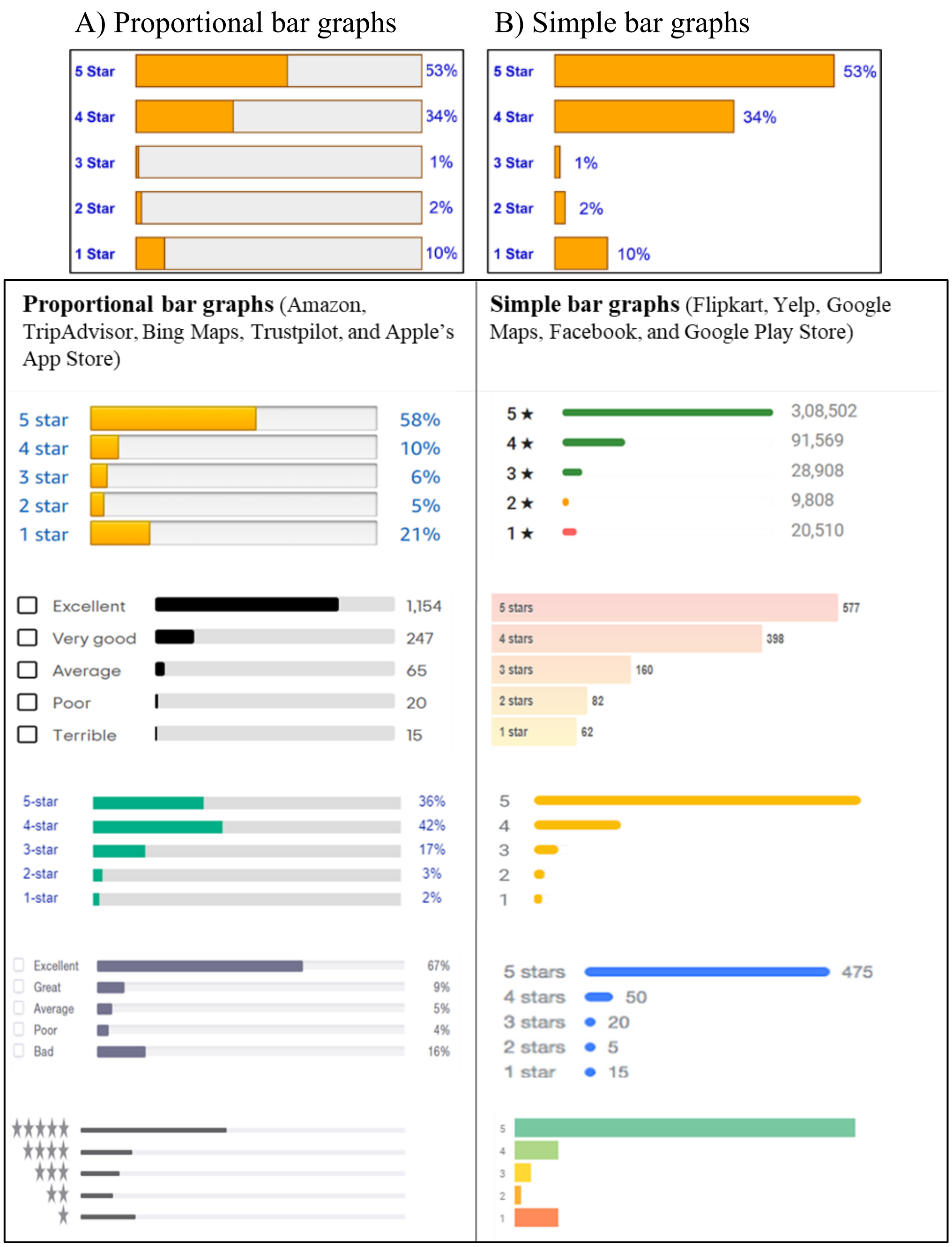
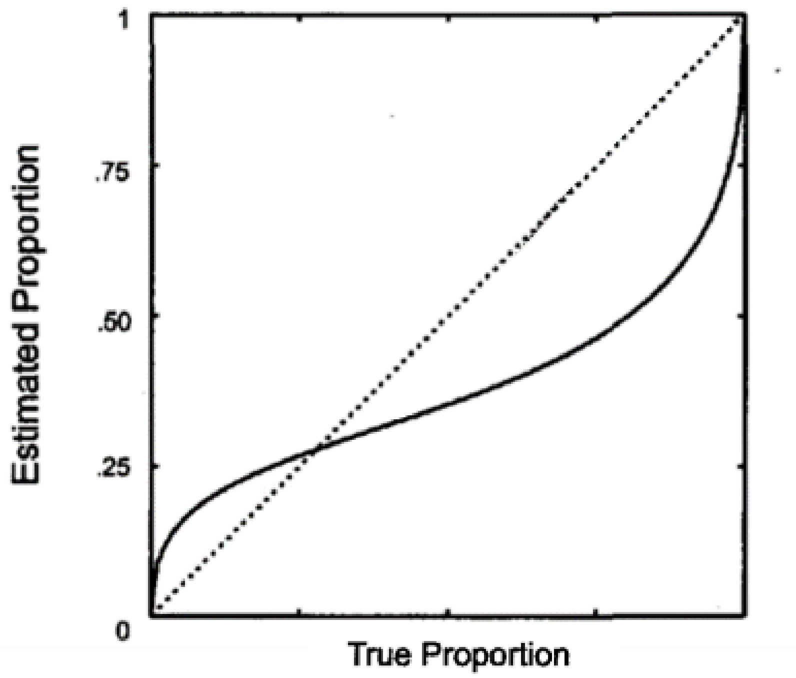
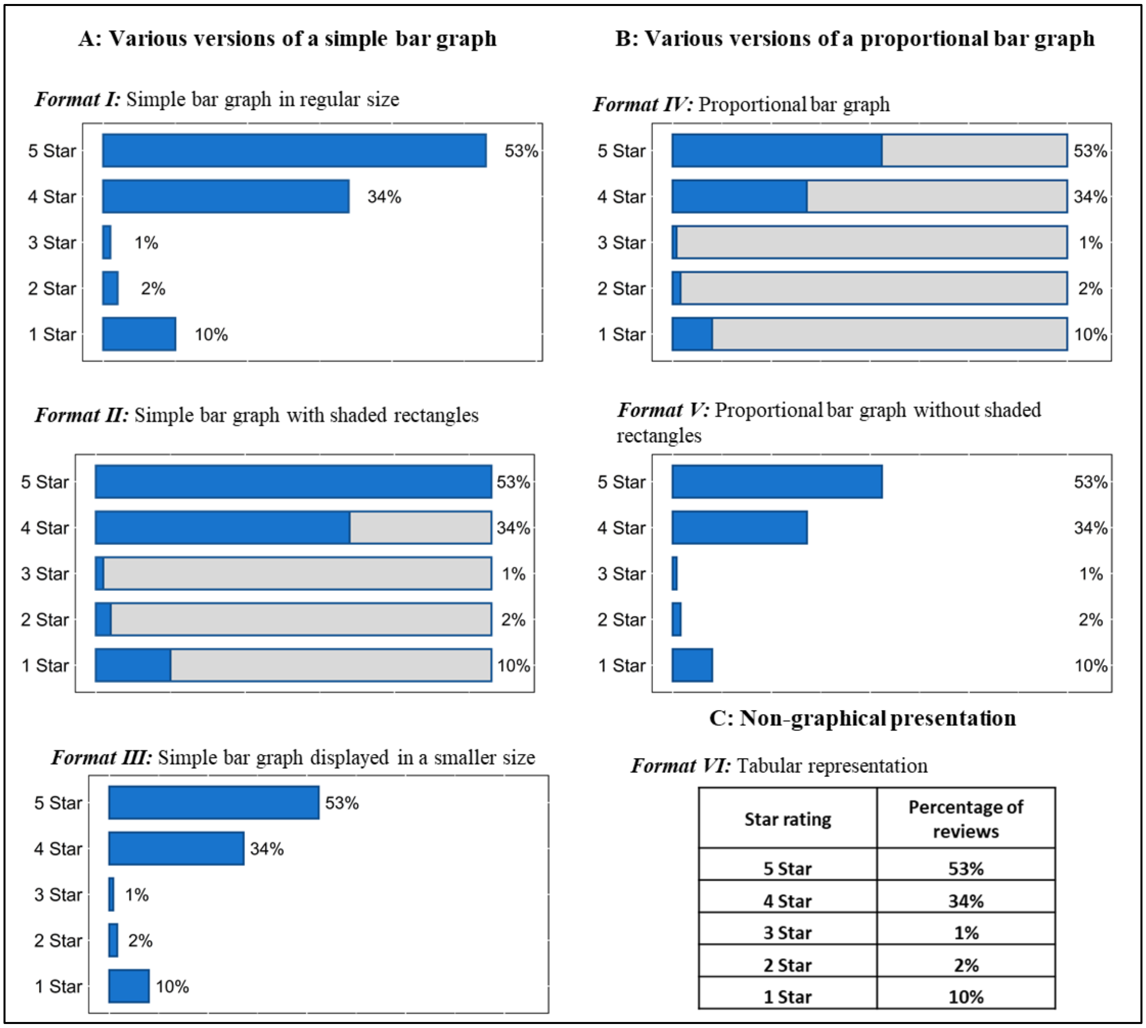


FIGURE 2: BIASES IN JUDGING PROPORTIONS PREDICTED BY TVERSKY AND KAHNEMAN (1992)



Source: (Hollands and Dyre 2000, p. 522)

FIGURE 3: EXAMPLES OF FORMATS USED IN STUDY 3



Star rating	Percentage of reviews
5 Star	53%
4 Star	34%
3 Star	1%
2 Star	2%
1 Star	10%

FIGURE 4: PRODUCT EVALUATION AS A FUNCTION OF GRAPHICAL FORMAT IN STUDY 3

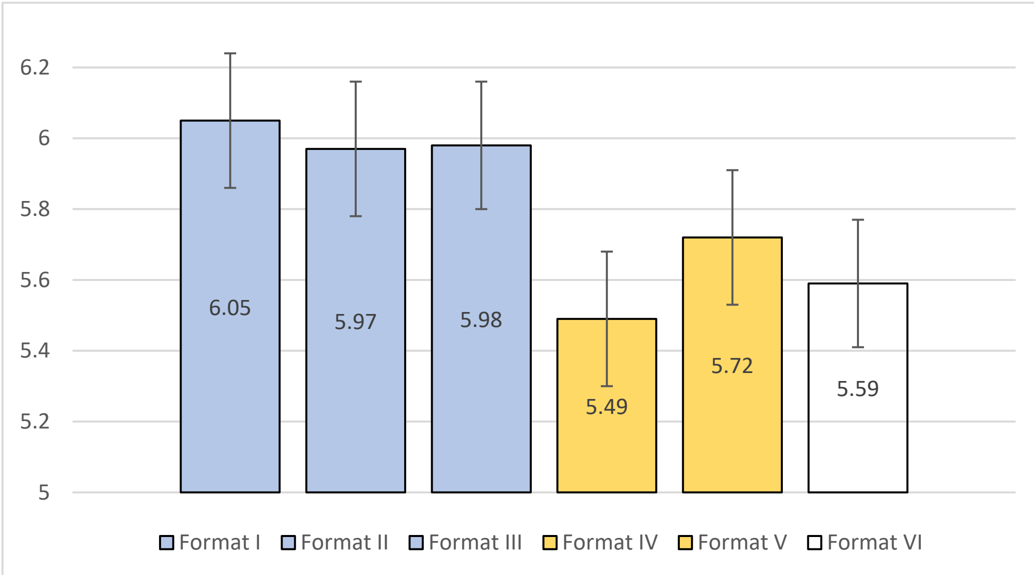
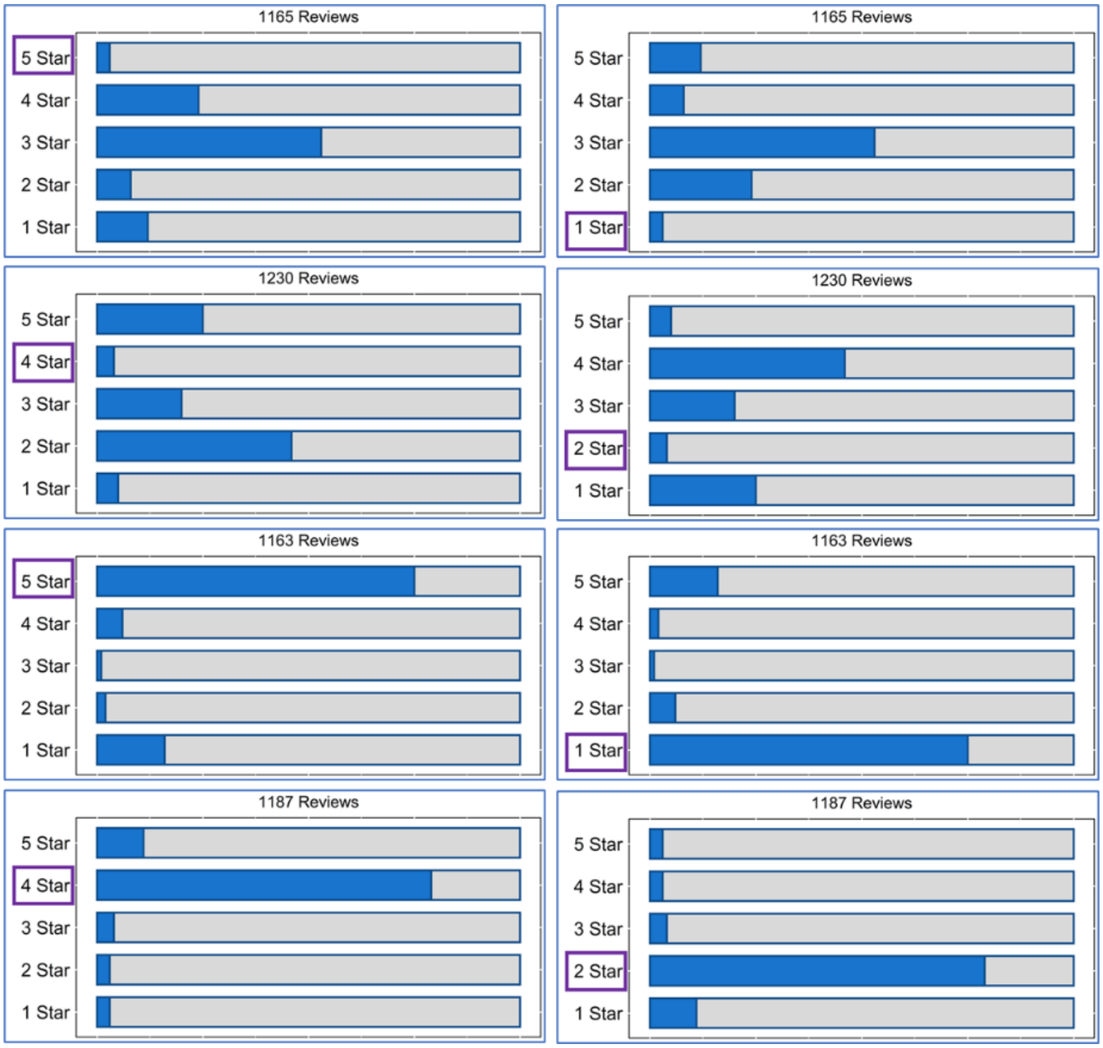


FIGURE 5: STIMULI AND TARGET RATING SCORES USED IN STUDY 4

A) Proportional bar graph (positive vs. negative condition)



B) Simple bar graphs (positive vs. negative condition)

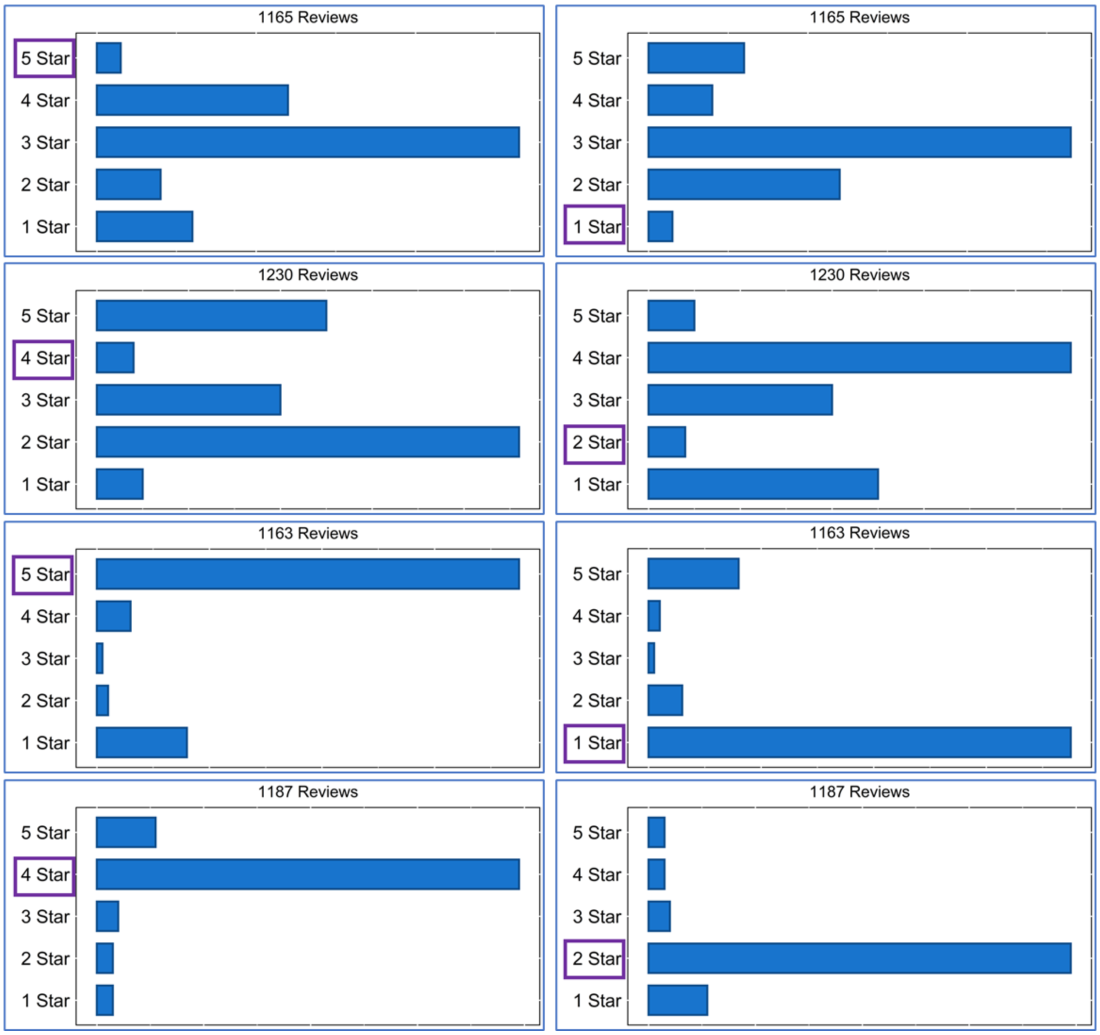


FIGURE 6: DISTRIBUTIONS AND STIMULI USED IN STUDY 5

Choice set A: Simple bar graphs | Choice set B: Proportional bar graphs

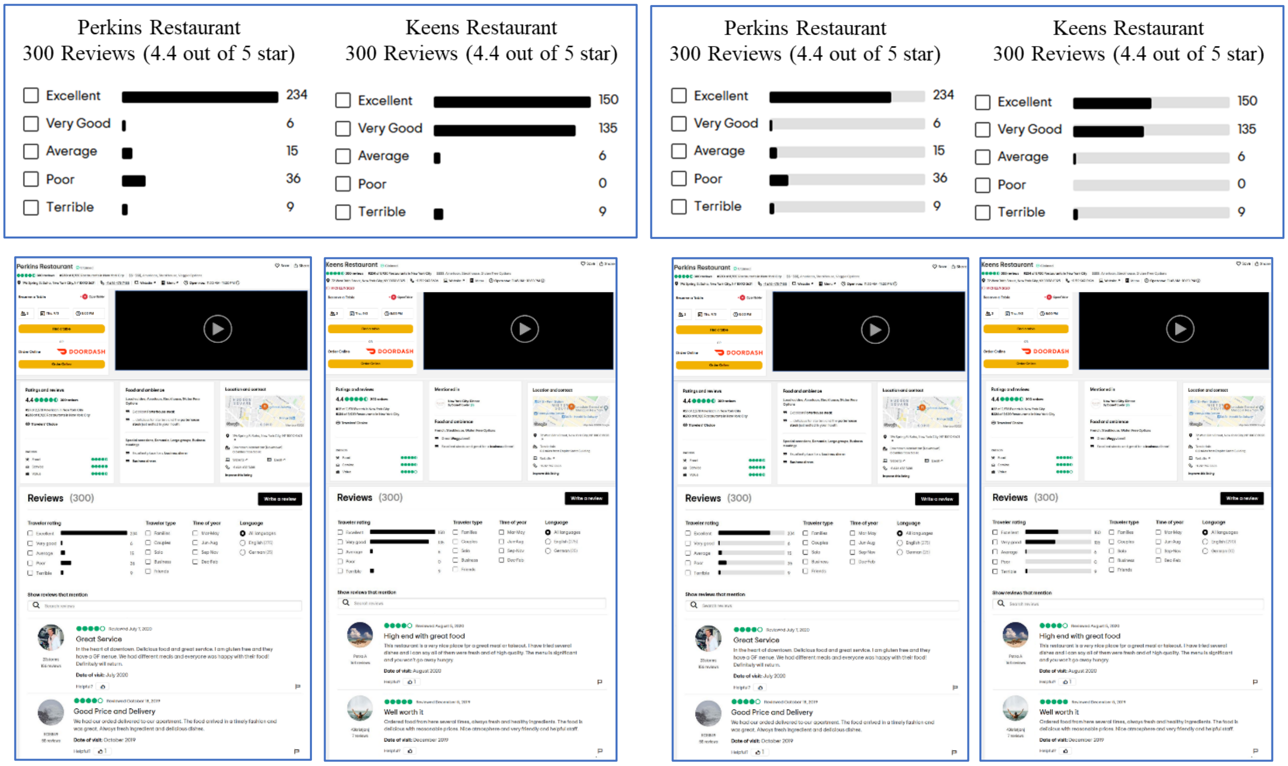
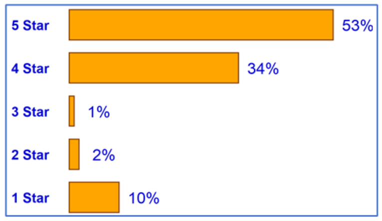
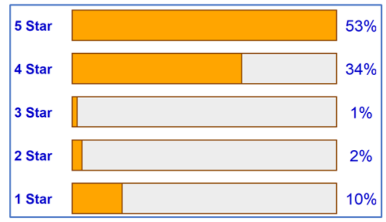


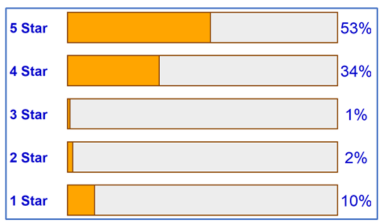
FIGURE 7: STIMULI USED IN APPENDIX A



Format I: Simple bar graph in regular size



Format II: Simple bar graph with blank spaces replaced by lightly shaded rectangles



Format III: Proportional bar graph with lightly shaded rectangles

FIGURE 8: WILLINGNESS TO BUY AS A FUNCTION OF THE GRAPHICAL FORMAT AND TIME GAZED AT GRAPHICAL CUES IN APPENDIX A

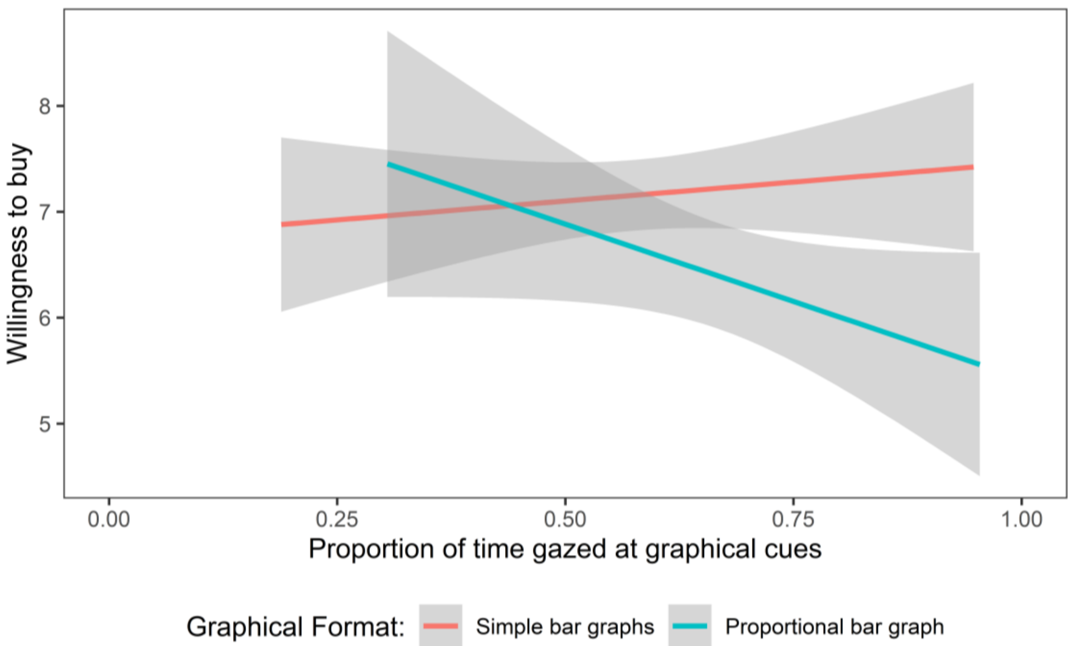
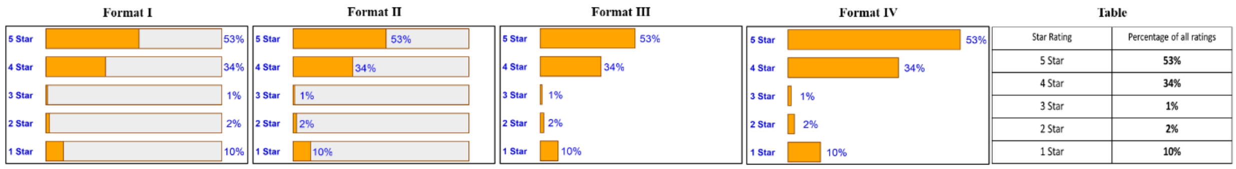


FIGURE 9: APPENDIX B FORMATS AND RESULTS

A) Formats used in Appendix B.



B) Appendix B results: Product evaluation as a function of rating distribution format.

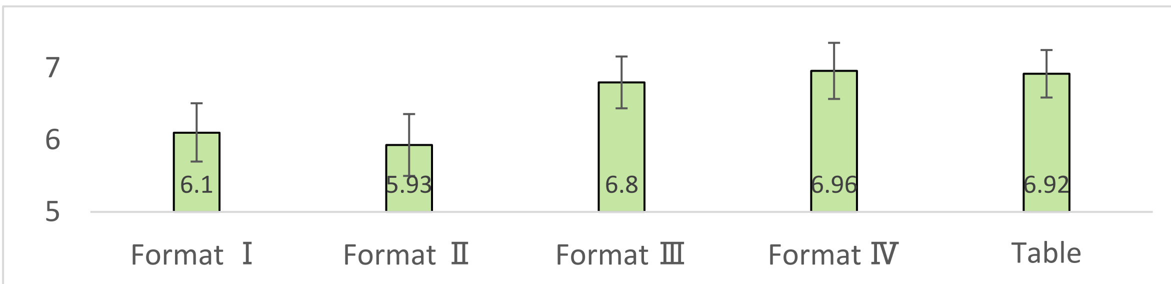


FIGURE 10: STIMULI USED IN APPENDIX C

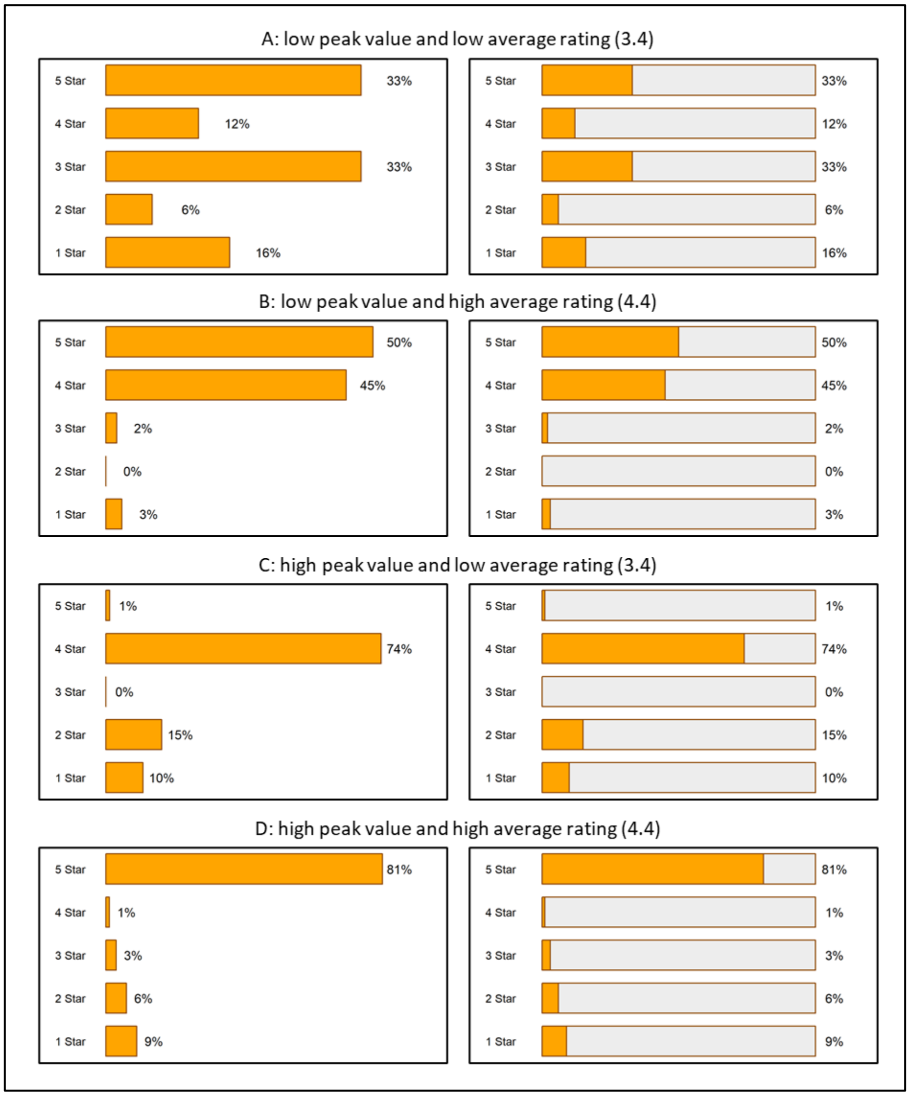
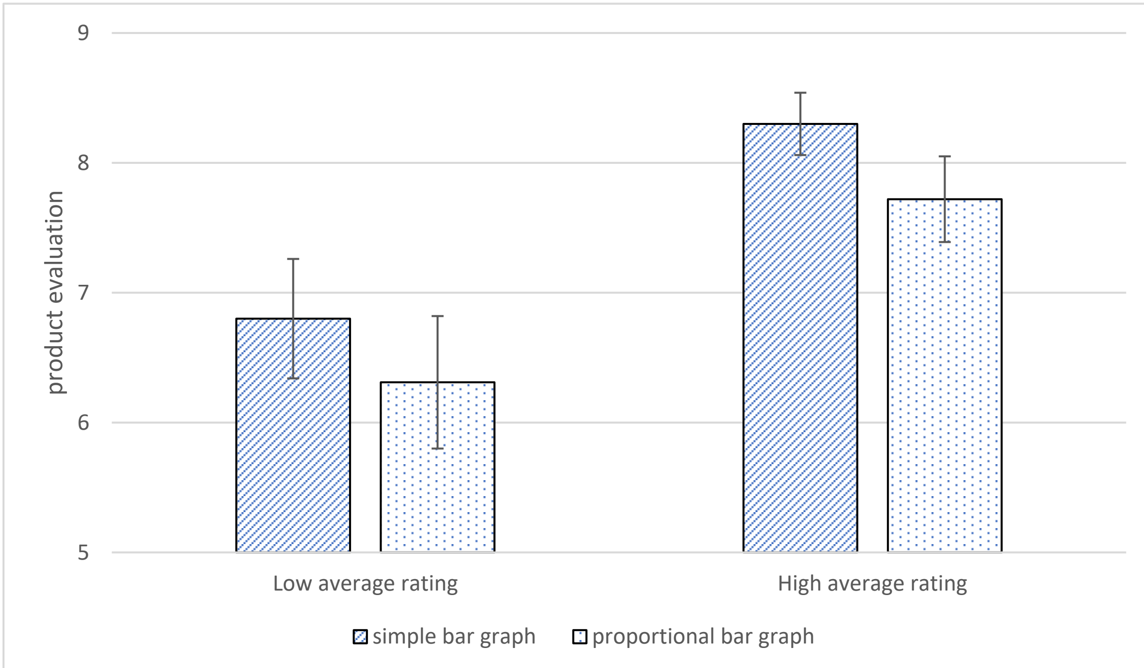


FIGURE 11: APPENDIX C RESULTS

A) Low peak value



B) High peak value

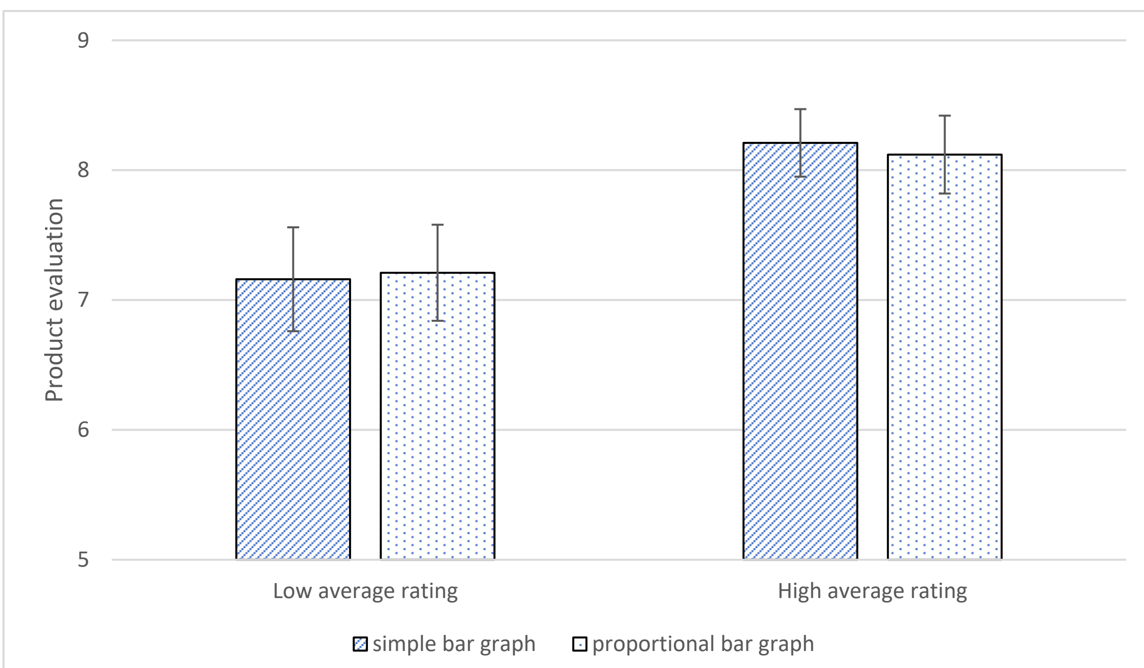
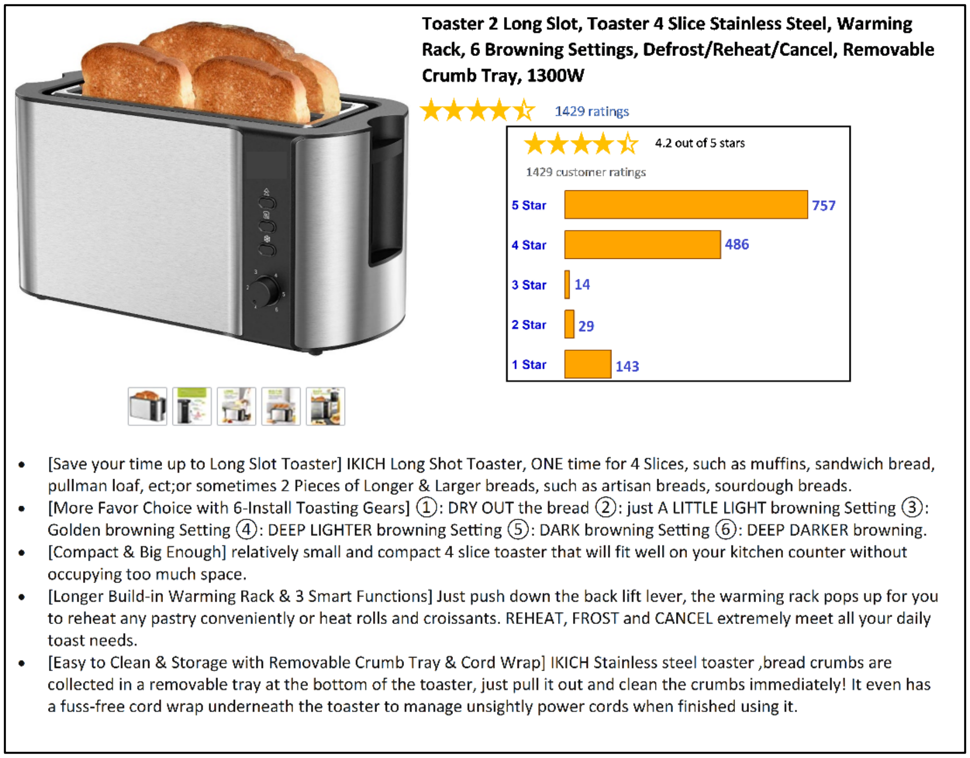
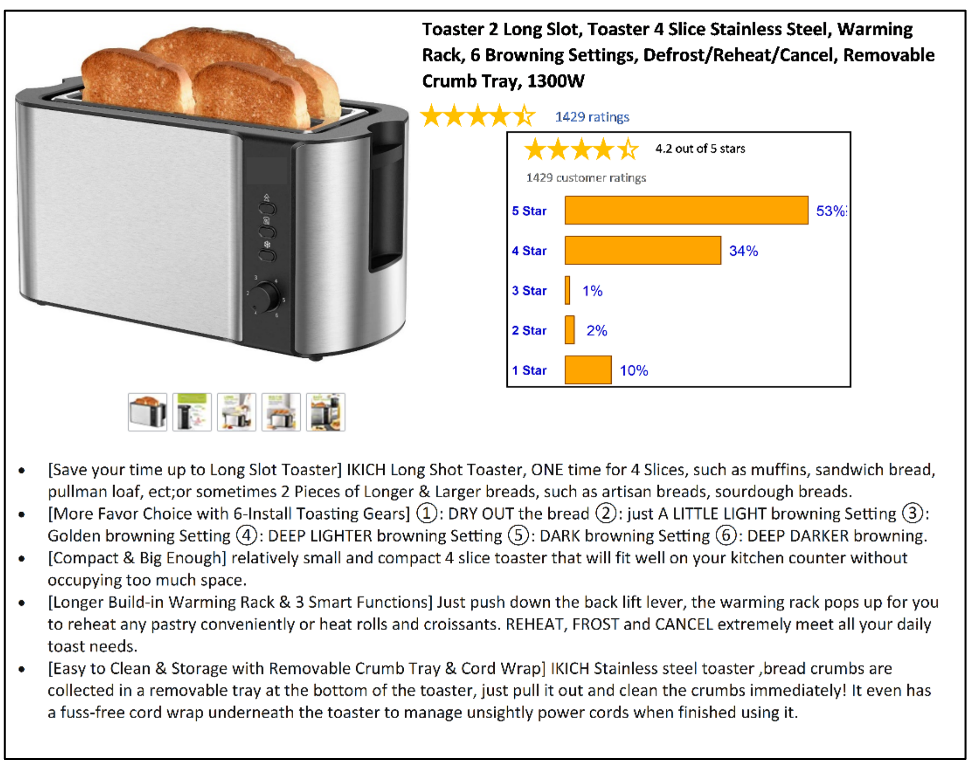


FIGURE 12: EXAMPLES OF STIMULI USED IN APPENDIX D

Format I (simple bar graph with frequencies)



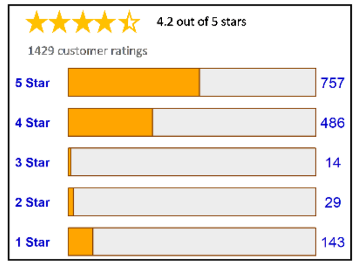
Format II (simple bar graph with percentages)



Format III (proportional bar graph with frequencies)

Toaster 2 Long Slot, Toaster 4 Slice Stainless Steel, Warming Rack, 6 Browning Settings, Defrost/Reheat/Cancel, Removable Crumb Tray, 1300W

1429 ratings



- [Save your time up to Long Slot Toaster] IKICH Long Shot Toaster, ONE time for 4 Slices, such as muffins, sandwich bread, pullman loaf, ect;or sometimes 2 Pieces of Longer & Larger breads, such as artisan breads, sourdough breads.
- [More Favor Choice with 6-Install Toasting Gears] ①: DRY OUT the bread ②: just A LITTLE LIGHT browning Setting ③: Golden browning Setting ④: DEEP LIGHTER browning Setting ⑤: DARK browning Setting ⑥: DEEP DARKER browning.
- [Compact & Big Enough] relatively small and compact 4 slice toaster that will fit well on your kitchen counter without occupying too much space.
- [Longer Build-in Warming Rack & 3 Smart Functions] Just push down the back lift lever, the warming rack pops up for you to reheat any pastry conveniently or heat rolls and croissants. REHEAT, FROST and CANCEL extremely meet all your daily toast needs.
- [Easy to Clean & Storage with Removable Crumb Tray & Cord Wrap] IKICH Stainless steel toaster ,bread crumbs are collected in a removable tray at the bottom of the toaster, just pull it out and clean the crumbs immediately! It even has a fuss-free cord wrap underneath the toaster to manage unsightly power cords when finished using it.

Format IV (proportional bar graph with percentages)

Toaster 2 Long Slot, Toaster 4 Slice Stainless Steel, Warming Rack, 6 Browning Settings, Defrost/Reheat/Cancel, Removable Crumb Tray, 1300W

1429 ratings

- [Save your time up to Long Slot Toaster] IKICH Long Shot Toaster, ONE time for 4 Slices, such as muffins, sandwich bread, pullman loaf, ect;or sometimes 2 Pieces of Longer & Larger breads, such as artisan breads, sourdough breads.
- [More Favor Choice with 6-Install Toasting Gears] ①: DRY OUT the bread ②: just A LITTLE LIGHT browning Setting ③: Golden browning Setting ④: DEEP LIGHTER browning Setting ⑤: DARK browning Setting ⑥: DEEP DARKER browning.
- [Compact & Big Enough] relatively small and compact 4 slice toaster that will fit well on your kitchen counter without occupying too much space.
- [Longer Build-in Warming Rack & 3 Smart Functions] Just push down the back lift lever, the warming rack pops up for you to reheat any pastry conveniently or heat rolls and croissants. REHEAT, FROST and CANCEL extremely meet all your daily toast needs.
- [Easy to Clean & Storage with Removable Crumb Tray & Cord Wrap] IKICH Stainless steel toaster ,bread crumbs are collected in a removable tray at the bottom of the toaster, just pull it out and clean the crumbs immediately! It even has a fuss-free cord wrap underneath the toaster to manage unsightly power cords when finished using it.

Format V (proportional bar graph with high contrast)

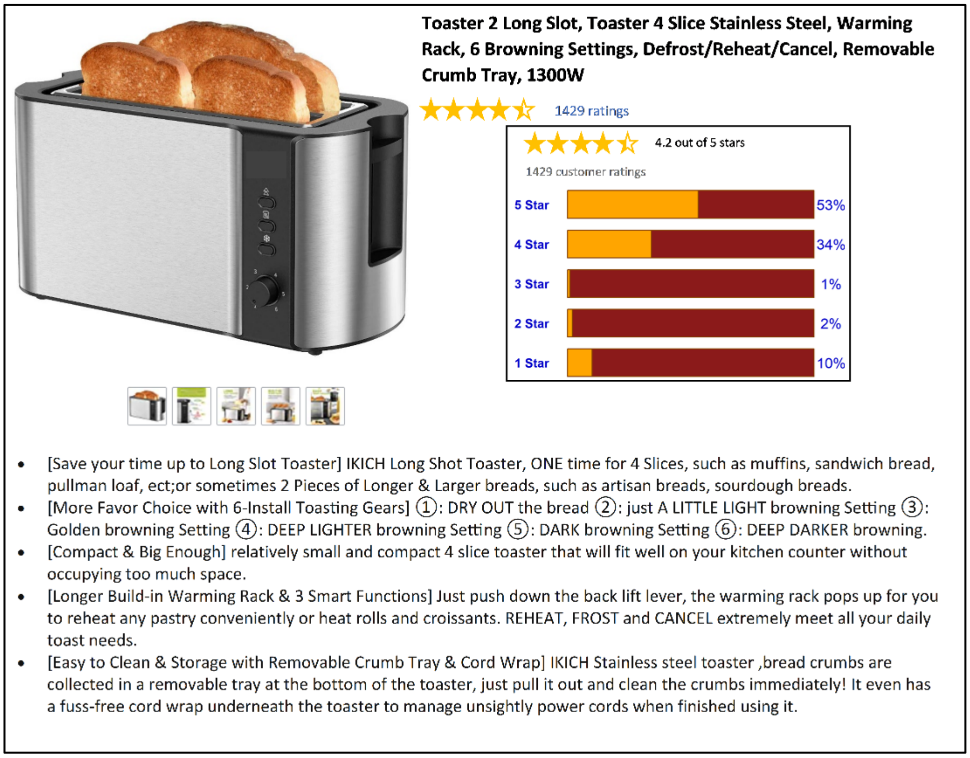
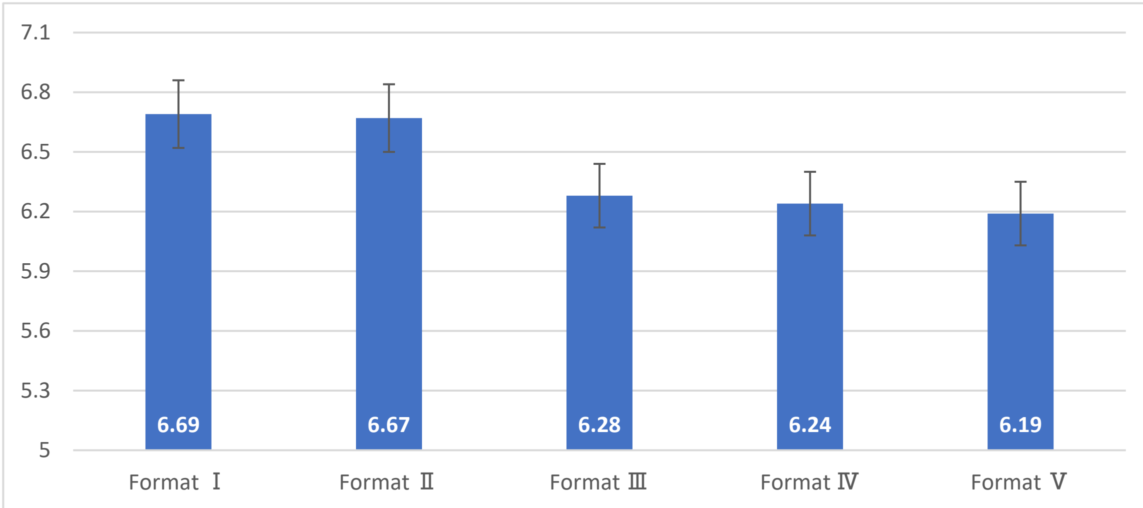


FIGURE 13: PRODUCT EVALUATIONS IN APPENDIX D



References

Ahluwalia, Rohini. 2002. “How Prevalent Is the Negativity Effect in Consumer Environments?” *Journal of Consumer Research* 29(2):270–79.

Bagozzi, Richard P. 2008. “Some Insights on Visual and Verbal Processing Strategies.” *Journal of Consumer Psychology* 18(4):258–63.

Barbot, Antoine, Sirui Liu, Ruth Kimchi, and Marisa Carrasco. 2017. “Attention Enhances Apparent Perceptual Organization.” *Psychonomic Bulletin and Review* 25(5):1824–32.

Baymard. 2017. “UX Research: 5 Requirements for the ‘Ratings Distribution Summary’ on the Product Page (65% Get It Wrong).” Retrieved (https://baymard.com/blog/user-ratings-distribution-summary).

Bettman, James R. and Pradeep Kakkar. 1977. “Effects of Information Presentation Format on Consumer Information Acquisition Strategies.” *Journal of Consumer Research* 3(4):233–40.

Bettman, James R., Mary Frances Luce, and John W. Payne. 1998. “Constructive Consumer Choice Processes.” *Journal of Consumer Research* 25(3):187–217.

Bonato, Mario, Sara Fabbri, Carlo Umiltà, and Marco Zorzi. 2007. “The Mental Representation of Numerical Fractions: Real or Integer?” *Journal of Experimental Psychology: Human Perception and Performance* 33(6):1410–19.

Chen, Zoey and Nicholas H. Lurie. 2013. “Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth.” *Journal of Marketing Research* 50(4):463–76.

Chevalier, Judith A. and Dina Mayzlin. 2006. “The Effect of Word of Mouth on Sales: Online Book Reviews.” *Journal of Marketing Research* 43(3):345–54.

Childers, Terry L., Michael J. Houston, and Susan E. Heckler. 1985. “Measurement of Individual Differences in Visual versus Verbal Information Processing.” *Journal of Consumer Research* 12(2):125–34.

Cian, Luca, Aradhna Krishna, and Ryan S. Elder. 2015. “A Sign of Things to Come: Behavioral Change through Dynamic Iconography.” *Journal of Consumer Research* 41(6):1426–46.

Cleveland, William S., Persi Diaconis, and Robert McGill. 1982. “Variables on Scatterplots Look More Highly Correlated When the Scales Are Increased.” *American Association for the Advancement of Science* 216(4550):1138–41.

Cleveland, William S. and Robert McGill. 1984. “Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods.” *Journal of the American Statistical Association* 79(387):531–54.

Cleveland, William S. and Robert McGill. 1986. "An Experiment in Graphical Perception." *International Journal of Man-Machine Studies* 25(5):491–500.

Cleveland, William S. and Robert McGill. 1987. "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data." *Journal of the Royal Statistical Society: Series A (General)* 150(3):192–210.

FinanceOnline. 2022. "102 Restaurant Software Statistics You Can't Ignore: 2022 Data, Analysis, & Market Share." Retrieved March 2, 2022 (https://financesonline.com/restaurant-software-statistics/).

Fischer, Martin H. 2000. "Do Irrelevant Depth Cues Affect the Comprehension of Bar Graphs?" *Applied Cognitive Psychology* 14(2):151–62.

Fisher, Matthew, George E. Newman, and Ravi Dhar. 2018. "Seeing Stars: How the Binary Bias Distorts the Interpretation of Customer Ratings." *Journal of Consumer Research* 45(3):471–89.

Gonzalez, Richard and George Wu. 1999. "On the Shape of the Probability Weighting Function." *Cognitive Psychology* 38:129–66.

Greenhouse, Samuel W. and Seymour Geisser. 1959. "On Methods in the Analysis of Profile Data." *Psychometrika* 24(2):95–112.

Grewal, Lauren and Andrew T. Stephen. 2019. "In Mobile We Trust: The Effects of Mobile Versus Nonmobile Reviews on Consumer Purchase Intentions." *Journal of Marketing Research* 56(5):791–808.

He, Stephen X. and Samuel D. Bond. 2015. "Why Is the Crowd Divided? Attribution for Dispersion in Online Word of Mouth." *Journal of Consumer Research* 41(6):1509–27.

Hegarty, Mary. 2011. "The Cognitive Science of Visual-Spatial Displays: Implications for Design." *Topics in Cognitive Science* 3(3):446–74.

Hollands, J. G. and Brian P. Dyre. 2000. "Bias in Proportion Judgments: The Cyclical Power Model." *Psychological Review* 107(3):500–524.

Hollands, J. G. and Ian Spence. 1992. "Judgments of Change and Proportion in Graphical Perception." *Human Factors* 34(3):313–34.

Hollands, J. G. and Ian Spence. 1998. "Judging Proportion with Graphs: The Summation Model." *Applied Cognitive Psychology* 12(2):173–90.

Hollands, J. G., Takeshi Tanaka, and Brian P. Dyre. 2002. "Understanding Bias in Proportion Production." *Journal of Experimental Psychology: Human Perception and Performance* 28(3):563–74.

Hu, Nan, Paul A. Pavlou, and Jie Zhang. 2017. "On Self-Selection Biases in Online Product Reviews." *MIS Quarterly* 41(2):449–71.

Judd, Charles M., Jacob Westfall, and David A. Kenny. 2012. "Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem." *Journal of Personality and Social Psychology* 103(1):54–69.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory." *Econometrica* 47(2):263–92.

Khare, Adwait, Lauren I. Labrecque, and Anthony K. Asare. 2011. "The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings." *Journal of Retailing* 87(1):111–26.

Kim, Junghan and Arun Lakshmanan. 2021. "Do Animated Line Graphs Increase Risk Inferences?" *Journal of Marketing Research*.

Klein, Jill G. and Rohini Ahluwalia. 2005. "Negativity in the Evaluation of Political Candidates." *Journal of Marketing* 69(1):131–42.

Kronrod, Ann and Shai Danziger. 2013. "'Wii Will Rock You!' The Use and Effect of Figurative Language in Consumer Reviews of Hedonic and Utilitarian Consumption." *Journal of Consumer Research* 40(4):726–39.

Landy, David, Brian Guay, and Tyler Marghetis. 2018. "Bias and Ignorance in Demographic Perception." *Psychonomic Bulletin and Review* 25(5):1606–18.

Lee, Angela Y. and Aparna A. Labroo. 2004. "The Effect of Conceptual and Perceptual Fluency on Brand Evaluation." *Journal of Marketing Research* 41(2):151–65.

Lurie, Nicholas H. and Charlotte H. Mason. 2007. "Visual Representation: Implications for Decision Making." *Journal of Marketing* 71(1):160–77.

Maslowska, Ewa, Edward C. Malthouse, and Stefan F. Bernritter. 2016. "Too Good To Be True: The Role of Online Reviews' Features in Probability to Buy." *International Journal of Advertising* 36(1):142–63.

Moe, Wendy W., Oded Netzer, and David A. Schweidel. 2017. "Social Media Analytics." Pp. 483–504 in *Handbook of Marketing Decision Models*, edited by B. Wierenga and R. Van der Lans. Springer International Publishing.

Moe, Wendy W. and David A. Schweidel. 2012. "Online Product Opinions: Incidence, Evaluation, and Evolution." *Marketing Science* 31(3):372–86.

Rocklage, Matthew D. and Russell H. Fazio. 2020. "The Enhancing Versus Backfiring Effects of Positive Emotion in Consumer Reviews." *Journal of Marketing Research* 57(2):332–52.

Rosario Babić, Ana, Francesca Sotgiu, Kristine De Valck, and Tammo H. A. Bijmolt. 2016. "The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors." *Journal of Marketing Research* 53(3):297–318.

Rozenkrants, Bella, S. Christian Wheeler, and Baba Shiv. 2017. "Self-Expression Cues in Product Rating Distributions: When People Prefer Polarizing Products." *Journal of Consumer Research* 44(4):759–77.

Scaife, Mike and Yvonne Rogers. 1996. "External Cognition : How Do Graphical Representations Work?" *International Journal of Human-Computer Studies* 45(2):1–29.

Schlosser, Ann E. 2005. "Posting versus Lurking: Communicating in a Multiple Audience Context." *Journal of Consumer Research* 32(2):260–65.

Schoenmueller, Verena, Oded Netzer, and Florian Stahl. 2020. "The Polarity of Online Reviews: Prevalence, Drivers and Implications." *Journal of Marketing Research* 57(5):853–77.

Schwarz, Norbert. 2004. "Metacognitive Experiences in Consumer Judgment and Decision Making." *Journal of Consumer Psychology* 14(4):332–48.

Shen, Hao, Yuwei Jiang, and Rashmi Adaval. 2010. "Contrast and Assimilation Effects of Processing Fluency." *Journal of Consumer Research* 36(5):876–89.

Slusser, Emily B., Rachel T. Santiago, and Hilary C. Barth. 2013. "Developmental Change in Numerical Estimation." *Journal of Experimental Psychology: General* 142(1):193–208.

Slusser, Emily and Hilary C. Barth. 2017. "Intuitive Proportion Judgment in Number-Line Estimation: Converging Evidence From Multiple Tasks." *Journal of Experimental Child Psychology* 162:181–98.

Spence, Ian and Peter Krizel. 1994. "Children's Perception of Proportion in Graphs." *Child Development* 65(4):1193–1213.

Stone, Eric R., J. Frank Yates, and Andrew M. Parker. 1997. "Effects of Numerical and Graphical Displays on Professed Risk-Taking Behavior." *Journal of Experimental Psychology: Applied* 3(4):243–56.

Sullivan, Jessica L., Barbara J. Juhasz, Timothy J. Slattery, and Hilary C. Barth. 2011. "Adults' Number-Line Estimation Strategies: Evidence from Eye Movements." *Psychonomic Bulletin and Review* 18(3):557–63.

Sun, Monic. 2012. "How Does the Variance of Product Ratings Matter?" *Management Science* 58(4):696–707.

Talbot, Justin, Vidya Setlur, and Anushka Anand. 2014. "Four Experiments on the Perception of Bar Charts." *IEEE Transactions on Visualization and Computer Graphics* 20(12):2152–60.

Thomas, Manoj and Ellie J. Kyung. 2019. “Slider Scale or Text Box: How Response Format Shapes Responses.” *Journal of Consumer Research* 45(6):1274–93.

Tversky, Amos and Daniel Kahneman. 1986. “Rational Choice and the Framing of Decisions.” *The Journal of Business* 59(4):251–278.

Tversky, Amos and Daniel Kahneman. 1992. “Advances in Prospect Theory : Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty* 5(4):297–323.

Varey, Carol A., Barbara A. Mellers, and Michael H. Birnbaum. 1990. “Judgments of Proportions.” *Journal of Experimental Psychology: Human Perception and Performance* 16(3):613–25.

Wang, Feng, Xuefeng Liu, and Eric Fang. 2015. “User Reviews Variance, Critic Reviews Variance, and Product Sales: An Exploration of Customer Breadth and Depth Effects.” *Journal of Retailing* 91(3):372–89.

Watson, Jared, Anastasiya Pocheptsova Ghosh, and Michael Trusov. 2018. “Swayed by the Numbers: The Consequences of Displaying Product Review Attributes.” *Journal of Marketing* 82(6):109–31.

West, Patricia M. and Susan M. Broniarczyk. 1998. “Integrating Multiple Opinions: The Role of Aspiration Level on Consumer Response to Critic Consensus.” *Journal of Consumer Research* 25(1):38–51.

Westerman, Deanne L., Meredith Lanska, and Justin M. Olds. 2015. “The Effect of Processing Fluency on Impressions of Familiarity and Liking.” *Journal of Experimental Psychology: Learning Memory and Cognition* 41(2):426–38.

Zhang, Hang and Laurence T. Maloney. 2012. “Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action, and Cognition.” *Frontiers in Neuroscience* 6:1–14.

Zhang, Hang, Xiangjuan Ren, and Laurence T. Maloney. 2020. “The Bounded Rationality of Probability Distortion.” *Proceedings of the National Academy of Sciences* 117(36):22024–34.

Chapter 2: Essay II—Political Conservatism and Complaining About Dissatisfactory Personal Customer-Employee Interaction

Political Conservatism and Complaining About Dissatisfactory Personal Customer-Employee Interaction

Abstract

Political ideology plays a vital role in shaping a wide range of consumers' behaviors in the marketplace. Specifically, previous research has shown that political identity affects customer satisfaction and complaining behaviors. In this research, the authors demonstrate that consumers with different political ideologies are not equally likely to complain about various attributes of a product or service. Specifically, the authors use a unique database of online reviews in conjunction with a county-level indicator of political ideology (the U.S. presidential election results between 2000 and 2020) and a lab experiment to show that conservative consumers are more likely to complain about their interpersonal interactions with firms' employees during a product or service delivery. Given that political ideology is an easily observable variable, the present research will help managers customize their offerings and allocate their resources more effectively.

Keywords: online WOM, political ideology, interpersonal interactions, empathy

Introduction

Private and public complaints are common reactions to product or service failures and can influence a brand's current and prospective customers. With the rise of technology, it has never been more straightforward for consumers to express their dissatisfaction privately to the suppliers or publicly on social media and online review platforms. Millions of internet users can easily see customers' negative word of mouth (WOM) on online review platforms. While negative WOM on review platforms and social media networks might negatively influence firms' performance, studying them offers valuable opportunities to identify the shortcomings, understand the underlying issues, and address them. Therefore, academics and practitioners must understand what types of product and service failures may cause more negative WOM.

Research on customer satisfaction and consumer complaints has studied the impact of various variables on satisfaction and complaint likelihood, including political identity (Fernandes et al. 2022; Jung et al. 2017), self-efficacy, Machiavellianism (Shoham, Gavish, and Segev 2012), impulsivity, and self-monitoring (Sharma et al. 2010). A related research stream has focused explicitly on consumer characteristics that can affect negative WOM (Casidy et al. 2021; Dubois, Bonezzi, and De Angelis 2016; Zhang, Feick, and Mittal 2014). While the findings of these studies help managers understand why some groups of consumers are less satisfied or more likely to spread negative WOM, they do not offer guidelines for detecting product and service attributes whose failures result in a higher dissatisfaction and an increased risk of complaining among different groups of customers. Therefore, in order to obtain the full benefits of consumer feedback and reduce future negative WOM, firms need to understand how different groups of customers react differently to various negative experiences.

The customer satisfaction literature has identified two fundamental types of attributes that determine customer satisfaction: core attributes (i.e., what is delivered or the outcome) and peripheral attributes (i.e., the delivery process—how is the product or service delivered; Anderson, Pearo, and Widener 2008). Peripheral attributes can be further categorized into two general subgroups, physical (i.e., environmental and mechanical components) and interactional attributes (i.e., interpersonal interactions involved in the delivery process). Importantly, not all customers value different aspects of a service or a product equally, resulting in different satisfaction compositions. For instance, it has been found that the effect of the core and peripheral attributes on overall satisfaction is moderated by consumers' characteristics such as gender, age, income, and previous experiences (Anderson et al. 2008).

Social interactions between customers and employees are inevitable in the marketplace. Given that such interactions occur overwhelmingly more in service contexts, it is critical to understand how various consumers react differently to their dissatisfactory interactions with the employees. Therefore, the current research focuses on the role of political ideology in shaping consumers' assessment of and reaction to dissatisfactory personal interactions with service employees.

Several studies have investigated the role of political ideology in a wide range of behaviors, including brand choice (Khan, Misra, and Singh 2013; Varman and Belk 2009), variety-seeking (Fernandes and Mandel 2014), customer satisfaction (Fernandes et al. 2022), and complaining (Jung et al. 2017). We build on previous research findings on political psychology and consumer behavior and contribute to this literature by investigating the impact of political ideology on negative WOM about dissatisfactory customer-employee interactions. The choice of political identity as a factor to be studied with respect to consumer post-consumption behaviors is useful because managers can use publicly available data (e.g., voter registration and polling data) to adjust their marketing practices according to the political orientations of their customers.

We contend that conservatives are more likely than liberals to share negative WOM about interpersonal interactions involved in product and service deliveries. Specifically, we hypothesize that because conservative (vs. liberal) consumers tend to adhere to social norms and show less empathy toward others, they are more critical about their personal interactions with employees. We test these predictions using online consumer reviews posted on Google and conducting randomized lab experiments. Finally, we discuss the theoretical and managerial implications of this work and some avenues for future research.

Theoretical Background

Post-consumption Behaviors: Satisfaction, Complaining, and Negative WOM

Consumer complaining is defined as any form of "behavioral expression of dissatisfaction" (Peter McGraw, Warren, and Kan 2015, p. 1153), and it includes a wide range of responses from private and public negative word of mouth to lengthy legal proceedings (Crie 2003). Such complaints are directed toward firms, third parties (e.g., Better Business Bureau), friends, family, or strangers. Given the increasing importance of establishing long-term relationships with customers, firms should encourage dissatisfied customers to lodge more complaints (Fornell and Wernerfelt 1987). In fact, actively encouraging complaints and monitoring negative WOM shared on social media and review platforms are standard practices among many businesses (Bell and Luddington 2006; Peter McGraw et al. 2015).

Consumer complaining has received considerable attention in the marketing literature as well. This includes research on motivational factors influencing consumers' complaining responses. Individuals often complain in different forms, with a variety of purposes, including venting negative emotions (Robertson and Shaw 2009), starting conversations with other consumers (Peter McGraw et al. 2015), warning others about a negative experience (Ward and Ostrom 2006), and demanding redress (Chebat, Davidow, and Codjovi 2005).

However, not all consumers engage in complaining activities equally and in the same manner. Negative WOM and complaining likelihood are influenced by a wide array of consumers' psychological characteristics. For instance, complaining propensity is associated with impulsivity (Sharma et al. 2010), self-confidence (Jacoby and Jaccard 1981), external locus of control (Robertson and Shaw 2009), assertiveness (Sharma et al. 2010; Voorhees, Brady, and Horowitz 2006), and self-efficacy (Mckee, Simmers, and Licata 2006). At a broader level, many researchers have studied the influence of a variety of cultural and situational factors on consumers complaining behaviors (e.g., Liu, Furrer, and Sudharshan 2001; Mattila and Patterson 2004; Orsingher, Valentini, and de Angelis 2010; Oster 1980; Schoefer 2010). For example, situational factors such as price (Oster 1980), the likelihood of success (Robertson and Shaw 2009), perceived social justice (Voorhees and Brady 2005), and sources of dissatisfaction (Bitner, Booms, and Tetreault 1990) have been found to influence consumers' complaining behaviors. Studying cultural differences in complaining, Liu and McClure (2001) report that consumers in individualistic cultures (e.g., United States) are more likely to engage in public complaining and less likely to show private reactions to dissatisfaction, as compared with those in collectivistic cultures (e.g., South Korea).

Many personal and cultural characteristics that have been found to influence consumers' complaining behaviors are related to political ideology. However, mixed findings have been reported regarding the way consumers' post-consumption behaviors such as satisfaction and negative WOM are influenced by personal and cultural characteristics associated with conservatism. For instance, political psychology research has found that conservatism is positively [negatively] associated with conscientiousness [openness to new experiences] (Carney et al. 2008). Nevertheless, research has found that higher degrees of both conscientiousness and openness to new experiences lead to a higher complaining likelihood (Ekinci, Calderon, and Siala 2016; Harris and Mowen 2001).

Such mixed findings can be found elsewhere (e.g., Phau and Baird 2008; Swimberghe, Sharma, and Flurry 2009). Importantly, Jung et al. (2017) find that while conservative consumers are less likely to complain about financial products, they are more likely to complain about services. On the one hand, religious beliefs–which are associated with conservatism (Noël and Thérien 2008)–have been found to be associated with higher negative WOM in everyday service failure encounters (Casidy et al. 2021). In contrast, recent research suggests that since conservatives believe in free will and trust their own decisions, they are more satisfied with both products and services and less likely to post negative reviews on public platforms (Fernandes et al. 2022).

Despite growing research on the role of political identity in the marketplace, it is still unclear why conservatism and closely related constructs are associated with favorable post-consumption behaviors in some cases and negative reactions in others. While the extant studies help understand the influence of political beliefs on various post-consumption behaviors such as complaining behaviors and customer satisfaction, most investigate consumers' overall satisfaction with products and services as the main driver of complaints without distinguishing between various sources of dissatisfaction. We believe that distinguishing between various product and service failures in exploring the influence of political ideology on negative WOM and complaints is essential to both theory and practice. Therefore, we study the effects of consumers' political ideologies on post-consumption behaviors (satisfaction and negative WOM) while focusing on customer-employee personal interactions. We specifically study consumers' evaluations of their personal interactions during service encounters to better understand how conservatives and liberals react differently to dissatisfactory social interactions in the marketplace.

Conservatism, Empathy, and Social Interactions

Political ideology refers to "beliefs and principles that reflect a person's views on how society should be governed" (Kim, Park, and Dubois 2018, p. 135) and is usually measured on a spectrum ranging from liberal to conservative. Studying its role in consumer behavior is crucial because individuals' political ideologies have been found to affect their thoughts, feelings, attitudes, and behaviors (Jost et al. 2003; Jung et al. 2017). Research has found that liberal and conservative consumers deviate from one another in various aspects, including personality traits, motivational interests, interpersonal relationships, and personal values (Carney et al. 2008; Jung et al. 2017).

Among those differences, some are more influential in shaping social interactions. Specifically, the two that stand out are empathy and social norms. Empathy is a multidimensional construct that affects individuals' assessment of social interactions and is defined as "*a person's ability to sense another's thoughts, feelings, and experiences, to share the* *other's emotional experience, and to react to the observed experiences of another person*" (Wieseke, Geigenmüller, and Kraus 2012, p. 317). Prior research suggests that while conservatives value politeness more than liberals do (Hirsh et al. 2010), they tend to show lower levels of compassion and empathy toward others (Hasson et al. 2018; Hirsh et al. 2010; Hudson, Cikara, and Sidanius 2019; McCue and David Gopoian 2000; Pratto, Sidanius, and Levin 2006). A consequence of differences in empathy felt toward others is that while liberals tend to attribute external factors to people's hardships and express more sympathy for them, conservatives blame individuals for their situations and show less compassion (Hasson et al. 2018).

Conservatives also tend to place a higher emphasis on traditional values and social norms, and prior research suggests that they are more sensitive to violations of those values and norms (Skitka and Tetlock 1993). Consistently, several studies have found that conservatism is associated with less tolerance for other people (Lindner and Nosek 2009; Noël and Thérien 2008). Conservatives tend to punish violators of social norms because such violations can increase uncertainty (Jost et al. 2003). Other studies have also found an association between constructs related to conservatism and adherence to social norms, unforgiveness, and retaliatory behaviors. For example, research has reported both an association (Noël and Thérien 2008) and a reciprocal causal relationship (Patrikios 2008) between religiosity and conservatism, and religious individuals have been found to advocate adherence to social norms and spread negative WOM about service failures more than nonreligious people do (Casidy et al. 2021). Furthermore, research indicates that southern United States residents, who are mainly conservative, have higher norms and standards for politeness and are more prone to respond to even "trivial" encounters such as an insulting remark or a petty argument than those who live in the north (Cohen et al. 1996 and 1999). These findings suggest that customer conservatism, empathy, and tolerance for others can have significant implications for personal interactions occurring in the marketplace.

Employee-Customer Interactions and Customer Satisfaction

Given that social interactions are ubiquitous in the production and delivery of products and services, a large corpus of research has focused on how interactions between consumers and frontline personnel affect consumers' evaluations of products and services (Bitner et al. 1990; Brady, Voorhees, and Brusco 2012; Hartline and Ferrell 1996; Wieseke et al. 2012; Yim, Tse, and Chan 2008). Prior studies suggest that consumers' assessment of the quality of their personal interactions with employees is a significant predictor of their overall satisfaction with the service in various industries (Bitner et al. 1990).

Specifically, the level of empathy displayed during personal interactions between customers and employees has been found to affect consumers' perception of service quality (Berry, Seiders, and Grewal 1967; Wieseke et al. 2012). In fact, Wieseke, Geigenmüller, and Kraus (2012) report that more empathic people tend to assess an unsatisfactory service encounter less negatively because empathy enables adaptive and prosocial behaviors. This implies that empathic customers, who are more likely to forgive mistakes occurring in service encounters, should also be less likely to spread negative WOM about their social interactions with employees.

Empathy involves seeing the world from another person's perspective and is characterized by feelings of compassion towards them (Hasson et al. 2018). By taking another person's perspective, an empathic individual can increase her understanding of other people's experiences and react appropriately to various circumstances (Wieseke et al. 2012). A better understanding of what is happening enables empathic customers to predict other people's behaviors, revise their expectations, and respond to them accordingly. In customer-employee interactions, customers with high empathy can similarly foster a more precise understanding of an employee's state of mind and adapt to the situation (Wieseke et al. 2012). The resulting adaptive approach improves a customer's social interactions with service employees, leading to higher satisfaction.

Another aspect of empathy involves feeling compassion for other people and being motivated to help others (Bendapudi, Singh, and Bendapudi 1996; Hasson et al. 2018; Pavey, Greitemeyer, and Sparks 2012). According to Berry, Seiders, and Grewal (1967), empathy is an other-focused emotion that reflects the need for harmony and motivates individuals to help others. As a result, an empathic customer may be compelled to be more mindful of employees' working conditions, recognize their efforts, and act less selfishly. Consistently, Wieseke, Geigenmüller, and Kraus (2012) find that empathic customers are less likely to engage in punitive behaviors when experiencing unsatisfying interactions with employees.

Considering all the evidence, given that conservative (vs. liberal) individuals value traditions, display less empathy toward others, advocate social norms, and show less tolerance for other people, they should be more critical of their dissatisfactory social interactions in the marketplace. Therefore, we predict that conservative consumers will evaluate their dissatisfactory personal interactions with the service personnel more negatively and complain more about those interactions. We further hypothesize that this effect is mediated by lower levels of empathy toward service employees felt by conservative consumers. In the following section, we report two studies to test the effect of political conservatism on the perception of interpersonal interactions in service contexts.

Study 1: Textual Content of Online Consumer Reviews

Using online reviews posted on review platforms such as Amazon, Yelp, and Google to understand consumers' post-consumption behaviors is common among researchers (e.g., Casidy et al. 2021; Fernandes et al. 2022). In this study, we use a unique dataset of reviews posted on Google to test whether conservative consumers are more likely to spread negative WOM about their dissatisfactory interactions with service employees.

We collected all reviews posted for 893 McDonald's branches on Google (data collection was done between January 7 and 9 of 2021. We selected this fast-food chain for several reasons. First, McDonald's is a well-known brand with many customers writing reviews for its' branches. Second, our analysis of Google data (i.e., Google Trends) did not suggest any significant difference in its popularity in liberal and conservative states in the U.S. Third, since all branches offer the same menu and similar services, the simplicity of shopping at this chain would minimize the number of attributes about which consumers would write reviews. All branches were located in the U.S. The following section describes the procedures used in this study.

Data Collection and Text Analysis

Data Collection. We selected 108 U.S. counties/parishes whose 2019 estimated populations were between 150,000 and 200,000. We selected this population range for several reasons. First and foremost, we could observe different levels of conservativism in those counties as measured by voting behaviors. Second, counties with higher [lower] populations were predominantly conservative [liberal], and including them in our sample would introduce unnecessary complexities and would make the data set biased. Third, we could assume that branches in those 108 counties (vs. those in counties with higher populations) were more likely to serve local customers than travelers. Finally, branches in those counties had received a considerable number of reviews.

After selecting the counties, we used 108 [*"McDonald's" in X county/parish, Y]* search phrases where X and Y denote county and state names, respectively. Then, we collected the reviews posted for the branches displayed on each of the 108 search result pages. This process resulted in 732,152 consumer ratings (352,099 written reviews) posted for 893 branches. Since search results were not exclusive to the target counties, several branches from neighbor counties were also collected and included in our analyses.

Preparing Text for Statistical Analysis. Analysis of the text in the reviews is complex for many reasons. First, there is no structure in the reviews. Most reviews written by consumers tend to be casual in their grammar and word usage. Second, the textual content in these reviews must be cleaned to remove words that are not informative. Third, many words must be transmuted so that they can be manipulated numerically. In the preprocessing step, the textual data is cleaned and standardized for analysis. We treated each individual review as a document and ran the following preprocessing steps. We removed non-English characters and words (e.g., HTML tags, URLs, telephone numbers, punctuation) that do not typically have any informational content about the writers' experiences. After preprocessing, each review was assumed to be an unordered collection of words with meaning. We also used the Quanteda R package to stem the words (i.e., convert words to their root forms—e.g., "write" for "writer," "wrote" and "writing") and remove all stop words that are not required for meaning (e.g., "and," "the," "as," "are," "by," "on," and "in").

Topic Modeling. Topic modeling is an unsupervised machine learning technique to discover topics in a text document automatically. Latent Dirichlet Allocation (LDA; Blei, Ng, and Jordan 2003) is one of the most commonly used topic modeling techniques and is capable of the extraction of a parsimonious set of an optimum number of topics (Tirunillai and Tellis 2014). We identified five popular topics for the reviews posted for various McDonald's branches on Google from the LDA analysis. In addition to identifying topics present in the text, the LDA algorithm also calculates the probability of each topic in individual reviews. For instance, if the probability of a review to include a topic is 70%, that means that 70% of the review's content is related to that topic.

We then used the most frequent keywords that appeared in each topic (extracted by the LDA algorithm) and the logical connection among them to name the topics. For instance, the topic name "Interpersonal Interactions" in Figure 1 was based on the words "manager/management," "employee," "customer," "rude," "worker," "back," "people," "call," and "lady." The naming process was verified by closely examining the reviews with the highest probabilities of each topic. We also determined the valence of each topic by examining the correlation between its probability and the reviews' star-rating scores. Topics' names and valence were as follows: Topic 1's valence was negative as its probability was negatively correlated with the reviews' rating scores ($r = -.25$, $p < .001$), and it was named "*Poor Interpersonal Interactions.*" Topic 2's valence was positive ($r = .24$, $p < .001$), and it was named *"High Food Quality."* Topic 3's valence was negative ($r = -.30$, $p < .001$), and it was called *"Poor Ordering Process."* Topic 4's valence was positive ($r = .45$, $p < .001$), and it was called *"General Positive Experience."* Finally, Topic 5 was of negative valence ($r = -.15$, $p < .001$), and it was called *"Poor Food Quality."*

[INSERT FIGURE 1 ABOUT HERE]

Hypothesis Testing

We predicted that conservative consumers were more likely to complain about their interactions with McDonald's staff in their reviews. At a branch level, each topic's likelihood was calculated by averaging the probability of that topic in each branch's reviews. To measure the impact of consumer conservatism on the probability of each topic for each branch, we used the average of U.S. presidential election results between 2000 and 2016. Following Jung et al. (2017), we computed a single index for political conservatism by subtracting the percentage of Democratic voters from the percentage of Republican voters and then averaging them in each county for the 2000, 2004, 2008, 2012, and 2016 elections. In five multiple-regression models, we regressed each branch level topic's probability on our conservatism index along with a number of county-level demographic variables and branch-level variables. Table 1 reports the regression models and model-free tests (i.e., correlation) of the relationship between the topics' probabilities and consumer conservatism.

Results. As reported in Table 1, we found that consumers who lived in more conservative communities were more likely to complain about their interpersonal interactions with the branch workers (Model 1: $b = .11$, SE = .03, $p < .001$). Our model-free test also supported the positive relationships between conservatism and the likelihood of complaining about shopping trips' interpersonal aspects ($r = .09$, $p = .007$). On the other hand, consumers in the conservative communities were less likely to appraise the food quality (Model 2: $\beta = –.08$, SE = .03, $p = .007$) and to share their positive experiences (Model 4: $\beta = –.04$, SE = .02, $p < .067$). Our model-free tests also supported such relationships ($r = -.12$, $p < .001$ and $r = -.11$, $p = .001$, respectively). Finally, conservatism did not predict either the likelihood of complaining about ordering process (Model 3: $\beta = .02$, SE = .03, $p = .401$), or the likelihood of complaining about the product (i.e., food) quality (Model 5: $\beta = -.01$, SE = .04, $p = .888$).

To summarize, we demonstrated that consumers' political ideology correlates significantly with the content of online reviews. Our analyses revealed that the effect of political ideology on topic probabilities is robust even after controlling for multiple country-level covariates. We found that while conservative consumers were more likely to complain about their interactions with service staff, they were not more sensitive to the food quality. Moreover, we found that conservative consumers were less likely to praise the food quality or talk about their positive experiences.

[INSERT TABLE 1 ABOUT HERE]

Study 2: Dissatisfactory Interactions with Service Employees: A Lab Experiment

Analyzing the field data in Study 1 provided evidence for our main hypothesis that conservative (vs. liberal) consumers are more likely to spread negative WOM about their personal interactions with employees. However, given the complexities of field data and many uncontrolled factors, testing our hypothesis in a more controlled setting was required. Therefore, Study 2 aimed to test the effect of political ideology on complaining about customer-employee personal interactions in a controlled scenario-based experiment. We conducted this experiment based on a scenario because scenarios make expensive and complex manipulations feasible, provide more control over uncontrolled variables, decrease random noise, and save time by summarizing the unfolding events (Bitner 1990).

Method. Three hundred fifty-four Prolific users (all native English speakers living in the U.S.) completed this study, out of whom 287 (143 females, M_{age} = 36.26 years) could pass the comprehension test, i.e., they could identify the key points described in the scenario. The sampling process was designed to have a balanced pool in terms of political belief: conservative vs. liberal.

Participants were randomly assigned to the product or interaction failure condition in a between-subjects design. In both conditions, participants were asked to imagine a shopping experience at a restaurant. In the product failure condition, participants were presented with a scenario in which the food quality was not acceptable, but the waiter treated them professionally and respectfully. On the other hand, in the interaction failure condition, they read a scenario in which the food quality was acceptable, but their interactions with the employee were dissatisfactory. Next, participants in both conditions were asked to report their tip amount (the percentage of the total bill) and willingness to complain to the restaurant manager, followed by a rating task, giving the restaurant ratings for service, food, and an overall rating. Finally, we measured participants' self-reported conservatism using a single item question anchored by 1 (very liberal) and 7 (very conservative) and some other basic demographic variables.

Manipulation check. We found that participants' evaluations of the service quality in the interaction failure condition (M = 1.30) was significantly lower than that of those in the low food quality condition (M = 3.98, $t(285) = –22.70$, $p < .001$; Cohen's d = 2.69). On the other hand, participants' evaluations of food quality in the interaction failure condition (M = 3.72) was significantly higher than that of those in the low food quality condition (M = 1.49, $t(285) = 20.65$, $p < .001$; Cohen's d = 2.45).

Results. To test the effect of conservatism on complaining likelihood, we regressed the complaining likelihood on failure type (interaction vs. product), conservatism, and their interaction. We found that failure type had a significant effect on the complaining likelihood, such that participants in the interaction-failure condition were less likely to complain about their experience (b = –2.26, S.E. = .53, p < .001). The main effect of conservatism was not significant (b = .12, S.E. = .08, p = .140), suggesting that conservatism did not affect the complaining likelihood about low food quality. However, the positive and significant interaction between conservatism and failure type (b = .30, S.E. = .12, p = .011) demonstrated that conservatism led to a higher complaining likelihood among participants in the interaction-failure condition.

We conducted the same regression analysis for participants' overall ratings that they would provide on a review platform. The main effect of conservatism on the overall rating was not significant (b = –.02, S.E. = .04, p = .558), suggesting that conservatism did not influence the overall rating in the product-failure condition. The positive main effect of failure type revealed that participants in the interaction-failure condition rated their experience more positively than those in the product-failure condition (b = .67, S.E. = .23, p = .004). Consistent with our prediction, however, we found a negative and significant interaction between the failure type and conservatism (b = –.11, S.E. = .05, p = .030), showing that conservative consumers in the interaction-failure condition gave fewer stars to the restaurant (Table 2).

[INSERT TABLE 2 ABOUT HERE]

General Discussion

This research investigates how consumers' political identities affect their post-consumption reactions to dissatisfactory service encounters. Although recent research has examined the role of consumers' political ideologies in forming various post-consumption behaviors such as customer satisfaction, word of mouth, and complaints (Fernandes et al. 2022; Jung et al. 2017; Jung and Mittal 2020), it is still unclear whether conservative consumers are more satisfied with all aspects of products and services. Given that personal interactions between customers and employees are ubiquitous in both product and service deliveries, understanding how conservative and liberal consumers react differently to dissatisfactory interactions with employees is theoretically and managerially critical.

Extending prior research on the role of political ideology in shaping consumers' behavior in the marketplace (Chan and Ilicic 2019; Fernandes et al. 2022; Fernandes and Mandel 2014; Jung et al. 2017; Jung and Mittal 2020; Ordabayeva and Fernandes 2018), we demonstrate the systematic effect of political identity on negative WOM and consumer complaining behavior. While prior research has found that conservative consumers are more satisfied with products and services and, therefore, less likely to post negative online reviews (Fernandes et al. 2022; Jung et al. 2017), we find that this is not always the case. Specifically, our results suggest that conservatives are more likely to complain about dissatisfactory interactions with employees.

This research also makes significant contributions to customer satisfaction research. Prior research has found that customers tend to differ in the weight they place on various service attributes (Anderson et al. 2008). Advancing this line of research, we find that compared to liberals, conservative consumers are more sensitive about the interpersonal aspects of their experience. Given an increasing focus on enhancing customer experience, our research draws specific attention to the differential impact of peripheral core attributes on consumers' satisfaction with their experiences.

These findings present intriguing avenues for future research. First, future studies could examine how political identity may affect consumers' perception of their interactions with employees over the Internet. Given that companies extensively use the Internet and social media to interact with their customers, whether consumers with various political beliefs respond to such interactions differently is both practically and theoretically important. It may be the case that the lack of social context in private chats over the internet may attenuate our observed effect. One example of this could be when consumers' requests are rejected by an online representative. However, for interactions happening on social media, conservatism may even result in more harsh reactions due to the public nature of social media interactions.

Second, more and more businesses are adopting virtual agents and chatbots to manage their interactions with consumers. Given that customer-virtual agent interactions are very similar to customer-employee interactions, future research could investigate how conservatives and liberals react to dissatisfactory interactions with such technologies. Such studies may lead businesses to take the political identity of their customers—which can be inferred from their I.P. address locations—into consideration in designing the artificial intelligence behind the chatbots and virtual agents.

Third, future research may investigate other boundary conditions for the observed effect. For example, when consumers observe negative social interactions between employees and other customers, differences in empathy for others may completely reverse the effect. In such cases, while lower empathy may lead consumers to ignore their peers, higher levels of empathy felt by liberals may result in lower evaluations of their experiences.

Our findings also have important implications for practitioners. First, given that political identity can be easily inferred from various sources such as polls and election records, companies could better predict how their customers respond to dissatisfactory interactions and train their frontline personnel according to their target segment.

Second, companies can use these results to assess the performance of their branches more objectively. Specifically, our results suggest that the same experience leads to different evaluations between conservative and liberal consumers. Therefore, higher volumes of negative WOM, or complaints about personal interactions in conservative areas should not necessarily be attributed to employees' incompetence.

Third, prior research has documented the reciprocal effects between customer satisfaction and employee satisfaction (Frey, Bayón, and Totzek 2013). High employee turnover rates can damage firms' profits. During the Covid-19 pandemic, hiring frontline employees and retaining them has become a more serious challenge for many businesses as people tend to quit their jobs more often. Enhancing employees' satisfaction can lower their turnover rates. Therefore, companies may use these results to improve employee satisfaction through higher customer satisfaction by evoking empathy in their conservative patrons by subtle interventions such as emphasizing similarities between employees and customers.

Tables

TABLE 1: STUDY 1 REGRESSION ANALYSES RESULTS

	Model 1: Topic 1 (poor Interactions)	Model 2: Topic 2 (high food quality)	Model 3: Topic 3 (poor ordering process)	Model 4: Topic 4 (general positive experience)	Model 5: Topic 5 (poor food quality)
Intercept*	.00 (.02, 1.000)	.00 (.02, 1.000)	.00 (.02, 1.000)	.00 (.02, 1.000)	.00 (.03, 1.000)
Conservatism	**.11 (.03, <.001)**	**–.08 (.03, .007)**	**.02 (.03, .401)**	**–.04 (.02, .067)**	**–.01 (.04, .888)**
Population	.02 (.03, .543)	.07 (.03, .020)	–.02 (.03, .463)	–.01 (.02, .506)	–.05 (.04, .203)
Poverty rate	.13 (.06, .022)	.03 (.06, .662)	–.21 (.05, <.001)	.11 (.04, .009)	–.10 (.07, .169)
Median income	–.01 (.06, .812)	.02 (.06, .758)	–.22 (.06, <.001)	.28 (.04, <.001)	–.21 (.08, .007)
Unemployment Rate	–.17 (.03, <.001)	–.17 (.03, <.001)	–.05 (.03, .085)	.35 (.02, <.001)	–.15 (.04, .001)
Education	.04 (.04, .285)	–.01 (.04, .845)	.03 (.04, .485)	.03 (.03, .365)	–.15 (.05, .008)
%Female	.00 (.03, .989)	.05 (.03, .131)	–.04 (.03, .161)	.02 (.02, .487)	–.03 (.04, .490)
%White	.29 (.13, .024)	–.52 (.13, <.001)	.02 (.12, .856)	–.12 (.1, .244)	.43 (.17, .013)
%Black	.40 (.11, <.001)	–.69 (.12, <.001)	.08 (.11, .469)	–.09 (.09, .308)	.35 (.15, .022)
%Asian	.10 (.08, .203)	–.17 (.09, .049)	.04 (.08, .648)	–.08 (.06, .193)	.17 (.11, .128)
Age (% of 65+)	.06 (.03, .044)	.05 (.03, .094)	–.10 (.03, <.001)	–.02 (.02, .460)	.05 (.04, .212)
Review characters	.45 (.03, <.001)	–.16 (.03, <.001)	.18 (.03, <.001)	–.39 (.02, <.001)	.06 (.04, .108)
Review counts	–.06 (.02, .027)	–.13 (.03, <.001)	.18 (.02, <.001)	.08 (.02, <.001)	–.17 (.03, <.001)
Store rating	–.27 (.03, <.001)	.40 (.03, <.001)	–.55 (.03, <.001)	.43 (.02, <.001)	–.11 (.04, .004)
R–squared	.523	.484	.554	.707	.130
Model–free test (r)	.09 (.007)	–.12 (<.001)	.10 (.003)	–.11 (.001)	.08 (.014)

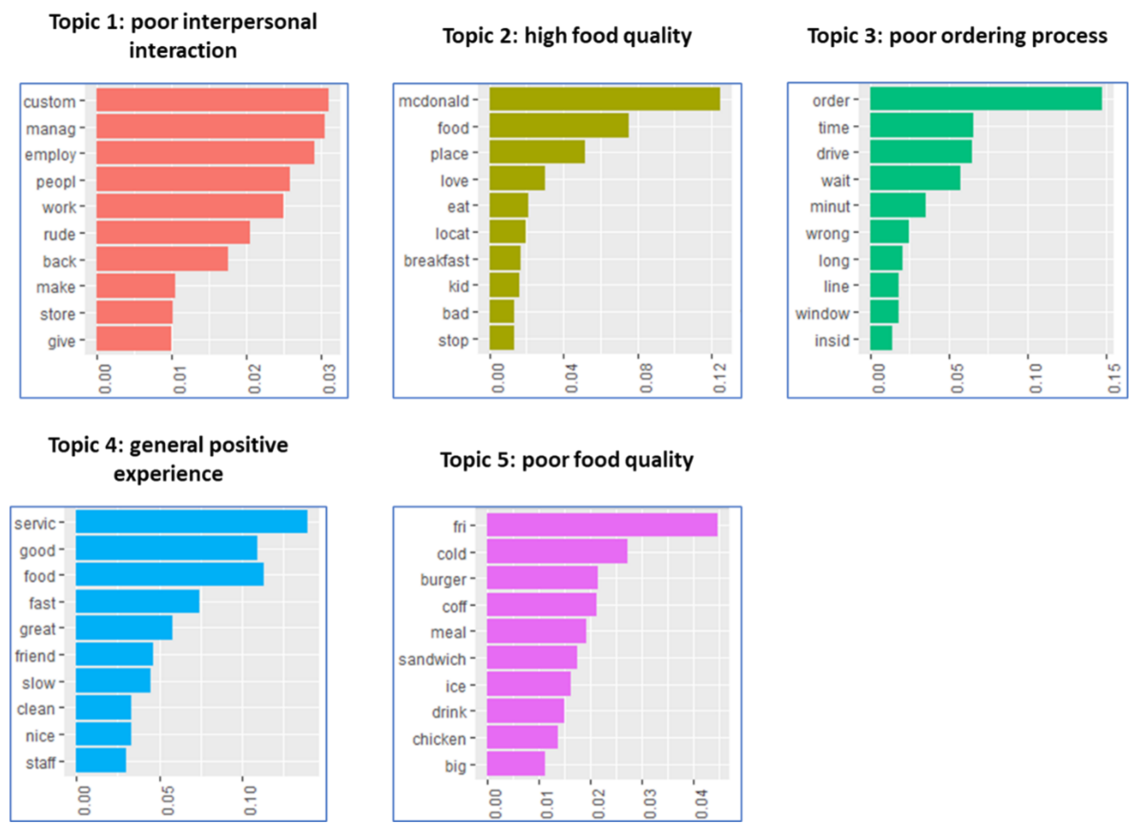
**β (SE, p-value)*
All variables are standardized.

TABLE 2: STUDY 2 REGRESSION ANALYSES RESULTS

	Complaining likelihood			Overall rating		
Predictor	*β*	*SE*	*p-value*	*β*	*SE*	*p-value*
Intercept	4.65	.37	< .001	2.45	.16	< .001
Failure type (interaction)	–2.26	.53	< .001	.67	.23	.004
Conservatism	.12	.08	.140	–.02	.04	.558
Failure type × Conservatism	.30	.12	.012	–.11	.05	.030
R-squared		.150			.062	

Figures

FIGURE 1: TOPICS' MOST IMPORTANT KEYWORDS FROM LDA ANALYS



References

Anderson, Shannon, Lisa Klein Pearo, and Sally K. Widener. 2008. "Drivers of Service Satisfaction: Linking Customer Satisfaction to the Service Concept and Customer Characteristics." *Journal of Service Research* 10(4):365–81.

Bell, Simon J. and James A. Luddington. 2006. "Coping With Customer Complaints." *Journal of Service Research* 8(3):221–33.

Bendapudi, Neeli, Surendra N. Singh, and Venkat Bendapudi. 1996. "Enhancing Helping Behavior: An Integrative Framework for Promotion Planning." *Journal of Marketing* 60(3):33–49.

Berry, Leonard L., Kathleen Seiders, and Dhruv Grewal. 1967. "Understanding Service Convenience." *Journal of Marketing* 66(3):1–17.

Bitner, Mary Jo. 1990. "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses." *Journal of Marketing* 54(2):69–82.

Bitner, Mary Jo, Bernard H. Booms, and Mary Stanfield Tetreault. 1990. "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents." *Journal of Marketing* 54(1):71–84.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3:993–1022.

Brady, Michael K., Clay M. Voorhees, and Michael J. Brusco. 2012. "Service Sweethearting: Its Antecedents and Customer Consequences." *Journal of Marketing* 76(2):81–98.

Carney, Dana R., John T. Jost, Samuel D. Gosling, and Jeff Potter. 2008. "The Secret Lives of Liberals and Conservatives: Personality Profiles, Interaction Styles, and the Things They Leave Behind." *Political Psychology* 29(6):807–40.

Casidy, Riza, Adam Duhachek, Vishal Singh, and Ali Tamaddoni. 2021. "Religious Belief, Religious Priming, and Negative Word of Mouth." *Journal of Marketing Research*.

Chan, Eugene Y. and Jasmina Ilicic. 2019. "Political Ideology and Brand Attachment." *International Journal of Research in Marketing* 36(4):630–46.

Chebat, Jean Charles, Moshe Davidow, and Isabelle Codjovi. 2005. "Silent Voices: Why Some Dissatisfied Consumers Fail to Complain." *Journal of Service Research* 7(4):328–42.

Cohen, Dov, Richard E. Nisbett, Brian F. Bowdle, and Norbert Schwarz. 1996. "Insult, Aggression, and the Southern Culture of Honor: An 'Experimental Ethnography.'" *Journal of Personality and Social Psychology* 70(5):945–60.

Cohen, Dov, Joseph Vandello, Sylvia Puente, and Adrian Rantilla. 1999. "'When You Call Me That, Smile!' How Norms for Politeness, Interaction Styles, and Aggression Work Together in Southern Culture." *Social Psychology Quarterly* 62(3):257–75.

Crie, Dominique. 2003. "Consumers' Complaint Behaviour. Taxonomy, Typology and Determinants: Towards a Unified Ontology." *Journal of Database Marketing & Customer Strategy Management* 11(1):60–79.

Dubois, David, Andrea Bonezzi, and Matteo De Angelis. 2016. "Sharing with Friends Versus Strangers: How Interpersonal Closeness Influences Word-of-Mouth Valence." *Journal of Marketing Research* 53(5):712–27.

Ekinci, Yuksel, Joana Calderon, and Haytham Siala. 2016. "Do Personality Traits Predict 'Complaining' Consumers?" *International Journal of Business Environment* 8(1):32–42.

Fernandes, Daniel and Naomi Mandel. 2014. "Political Conservatism and Variety-Seeking." *Journal of Consumer Psychology* 24(1):79–86.

Fernandes, Daniel, Nailya Ordabayeva, Kyuhong Han, Jihye Jung, and Vikas Mittal. 2022. "How Political Identity Shapes Customer Satisfaction." *Journal of Marketing* 1–19.

Fornell, Claes and Birger Wernerfelt. 1987. "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis." *Journal of Marketing Research* 24(4):337–46.

Frey, Regina Viola, Tomás Bayón, and Dirk Totzek. 2013. "How Customer Satisfaction Affects Employee Satisfaction and Retention in a Professional Services Context." *Journal of Service Research* 16(4):503–17.

Harris, Eric G. and John C. Mowen. 2001. "The Influence of Cardinal-, Central-, and Surface-Level Personality Traits on Consumers' Bargaining and Complaint Intentions." *Psychology and Marketing* 18(11):1155–85.

Hartline, Michael D. and O. C. Ferrell. 1996. "The Management of Customer-Contact Service Employees: An Empirical Investigation." *Journal of Marketing* 60(4):52–70.

Hasson, Yossi, Maya Tamir, Kea S. Brahms, J. Christopher Cohrs, and Eran Halperin. 2018. "Are Liberals and Conservatives Equally Motivated to Feel Empathy Toward Others?" *Personality and Social Psychology Bulletin* 44(10):1449–59.

Hirsh, Jacob B., Colin G. DeYoung, Xiaowen Xu, and Jordan B. Peterson. 2010. "Compassionate Liberals and Polite Conservatives: Associations of Agreeableness With Political Ideology and Moral Values." *Personality and Social Psychology Bulletin* 36(5):655–64.

Hudson, Sa kiera Tiarra Jolynn, Mina Cikara, and Jim Sidanius. 2019. "Preference for Hierarchy Is Associated with Reduced Empathy and Increased Counter-Empathy towards Others, Especially out-Group Targets." *Journal of Experimental Social Psychology* 85(August):103871.

Jacoby, Jacob and James J. Jaccard. 1981. "The Sources, Meaning, and Validity of Consumer Complaint Behavior: A Psychological Analysis." *Journal of Retailing* 57(3):4–24.

Jost, John T., Jack Glaser, Arie W. Kruglanski, and Frank J. Sulloway. 2003. "Political Conservatism as Motivated Social Cognition." *Psychological Bulletin* 129(3):339–75.

Jung, Jihye and Vikas Mittal. 2020. "Political Identity and the Consumer Journey: A Research Review." *Journal of Retailing* 96(1):55–73.

Jung, Kiju, Ellen Garbarino, Donnel A. Briley, and Jesse Wynhausen. 2017. "Blue and Red Voices: Effects of Political Ideology on Consumers' Complaining and Disputing Behavior." *Journal of Consumer Research* 44(3):477–99.

Khan, Romana, Kanishka Misra, and Vishal Singh. 2013. "Ideology and Brand Consumption." *Psychological Science* 24(3):326–33.

Kim, Jeehye Christine, Brian Park, and David Dubois. 2018. "How Consumers' Political Ideology and Status-Maintenance Goals Interact to Shape Their Desire for Luxury Goods." *Journal of Marketing* 82(6):132–49.

Lindner, Nicole M. and Brian A. Nosek. 2009. "Alienable Speech: Ideological Variations in the Application of Free-Speech Principles." *Political Psychology* 30(1):67–92.

Liu, Ben Shaw-ching, Olivier Furrer, and D. Sudharshan. 2001. "Culture and Behavioral Intentions Toward Services." *Journal of Service Research* 4(2):118–29.

Liu, Raymond R. and Peter McClure. 2001. "Recognizing Cross-Cultural Differences in Consumer Complaint Behavior and Intentions: An Empirical Examination." *Journal of Consumer Marketing* 18(1):54–74.

Mattila, Anna S. and Paul G. Patterson. 2004. "Service Recovery and Fairness Perceptions in Collectivist and Individualist Contexts." *Journal of Service Research* 6(4):336–46.

McCue, Clifford P. and J. David Gopoian. 2000. "Dispositional Empathy and the Political Gender Gap." *Women and Politics* 21(2):1–20.

Mckee, Daryl, Christina S. Simmers, and Jane Licata. 2006. "Customer Self-Efficacy and Response to Service." *Journal of Service Research* 8(3):207–20.

Noël, Alain and Jean-Philippe Thérien. 2008. *Left and Right in Global Politics*. Cambridge, UK: Cambridge University Press.

Ordabayeva, Nailya and Daniel Fernandes. 2018. "Better or Different? How Political Ideology

Shapes Preferences for Differentiation in the Social Hierarchy." *Journal of Consumer Research* 45(2):227–50.

Orsingher, Chiara, Sara Valentini, and Matteo de Angelis. 2010. "A Meta-Analysis of Satisfaction with Complaint Handling in Services." *Journal of the Academy of Marketing Science* 38(2):169–86.

Oster, Sharon. 1980. "The Determinants of Consumer Complaints." *The Review of Economics and Statistics* 62(4):603–9.

Patrikios, Stratos. 2008. "American Republican Religion? Disentangling the Causal Link Between Religion and Politics in the US." *Political Behavior* 30(3):367–89.

Pavey, Louisa, Tobias Greitemeyer, and Paul Sparks. 2012. "'I Help Because I Want to, Not Because You Tell Me to': Empathy Increases Autonomously Motivated Helping." *Personality and Social Psychology Bulletin* 38(5):681–89.

Peter McGraw, A., Caleb Warren, and Christina Kan. 2015. "Humorous Complaining." *Journal of Consumer Research* 41(5):1153–71.

Phau, Ian and Michael Baird. 2008. "Complainers Versus Non-Complainers Retaliatory Responses Towards Service Dissatisfactions." *Marketing Intelligence and Planning* 26(6):587–604.

Pratto, Felicia, Jim Sidanius, and Shana Levin. 2006. "Social Dominance Theory and the Dynamics of Intergroup Relations: Taking Stock and Looking Forward." *European Review of Social Psychology* 17(1):271–320.

Robertson, Nichola and Robin N. Shaw. 2009. "Predicting the Likelihood of Voiced Complaints in the Self-Service Technology Context." *Journal of Service Research* 12(1):100–116.

Schoefer, Klaus. 2010. "Cultural Moderation in the Formation of Recovery Satisfaction Judgments: A Cognitive-Affective Perspective." *Journal of Service Research* 13(1):52–66.

Sharma, Piyush, Roger Marshall, Peter Alan Reday, and Woonbong Na. 2010. "Complainers Versus Non-Complainers: A Multi-National Investigation of Individual and Situational Influences on Customer Complaint Behaviour." *Journal of Marketing Management* 26(1–2):163–80.

Shoham, Aviv, Yossi Gavish, and Sigal Segev. 2012. "Drivers of Customers' Reactions to Service Failures: The Israeli Experience." *International Journal of Psychological Studies* 4(1):76–90.

Skitka, Linda J. and Philip E. Tetlock. 1993. "Providing Public Assistance: Cognitive and Motivational Processes Underlying Liberal and Conservative Policy Preferences." *Journal of Personality and Social Psychology* 65(6):1205–23.

Swimberghe, Krist, Dheeraj Sharma, and Laura Flurry. 2009. “An Exploratory Investigation of the Consumer Religious Commitment and Its Influence on Store Loyalty and Consumer Complaint Intentions.” *Journal of Consumer Marketing* 26(5):340–47.

Tirunillai, Seshadri and Gerard J. Tellis. 2014. “Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation.” *Journal of Marketing Research* 2437:463–79.

Varman, Rohit and Russell W. Belk. 2009. “Nationalism and Ideology in an Anticonsumption Movement.” *Journal of Consumer Research* 36(4):686–700.

Voorhees, Clay M. and Michael K. Brady. 2005. “A Service Perspective on the Drivers of Complaint Intentions.” *Journal of Service Research* 8(2):192–204.

Voorhees, Clay M., Michael K. Brady, and David M. Horowitz. 2006. “A Voice From the Silent Masses: An Exploratory and Comparative Analysis of Noncomplainers.” *Journal of the Academy of Marketing Science* 34(4):514–27.

Ward, James C. and Amy L. Ostrom. 2006. “Complaining to the Masses: The Role of Protest Framing in Customer-Created Complaint Web Sites.” *Journal of Consumer Research* 33(2):220–30.

Wieseke, Jan, Anja Geigenmüller, and Florian Kraus. 2012. “On the Role of Empathy in Customer-Employee Interactions.” *Journal of Service Research* 15(3):316–31.

Yim, Chi Kin, David K. Tse, and Kimmy Wa Chan. 2008. “Strengthening Customer Loyalty through Intimacy and Passion: Roles of Customer—Firm Affection and Customer—Staff Relationships in Services.” *Journal of Marketing Research* 45(6):741–56.

Zhang, Yinlong, Lawrence Feick, and Vikas Mittal. 2014. “How Males and Females Differ in Their Likelihood of Transmitting Negative Word of Mouth.” *Journal of Consumer Research* 40(6):1097–1108.