The Polarization Effect of Ranked List Item Presentation on Item Evaluation

Authors: Duan Shen, Meng Lu, Meng Lu

Date: 2021-08-16T00:00:00+00:00

Abstract

Ranking lists serve as a tool for consumers to rapidly identify the relative quality of brands or products. Grounded in spatial metaphor theory, this study examines the polarizing effect of different ranking list presentation formats (vertical versus horizontal) on consumer evaluations of list items and the underlying mechanism. Through five experiments, we find that consumers exhibit greater evaluation polarization for items in vertical ranking lists compared to horizontal ranking lists, and we verify the mediating role of perceived differences among items in this relationship; that is, the evaluation polarization effect of vertical ranking lists is driven by individuals' higher perceived differences among ranking list items. Furthermore, the evaluation polarization effect across different list items only emerges when attribute evaluability is low; conversely, when the attribute evaluability of ranking list items is high, this evaluation polarization effect disappears.

Full Text

The Polarization Effect of Ranking List Presentation Formats on Item Evaluation

DUAN Shen, MENG Lu
(School of Business, Renmin University of China, Beijing 100872, China)

Abstract: Ranking lists serve as a communication tool to help consumers quickly identify the advantages and disadvantages of brands or products. Drawing on spatial metaphor theory, this research investigates how different presentation formats (vertical vs. horizontal) of ranking lists influence the polarization of consumer evaluations and the underlying mechanism. Through five experiments, we find that consumers exhibit greater evaluation polarization for items in vertical ranking lists compared to horizontal ranking lists. We further demonstrate that perceived item differences mediate this relationship—the polarization

effect in vertical ranking lists is driven by consumers' heightened perception of differences among ranked items. Moreover, this polarization effect emerges only when attribute evaluability is low; conversely, when attribute evaluability is high, the evaluation polarization effect disappears.

Keywords: ranking list, evaluation polarization effect, perceived item differences, attribute evaluability, spatial metaphor
Classification Code: B849: F713.55

Third-party entities such as *U.S. News & World Report* and *Bon Appetit* typically employ vertical ranking lists, as these rankings encompass multiple target items across broad categories (e.g., *QS World University Rankings*, *Fortune* 500) rather than limiting themselves to a few brands or single-product sorting. To accommodate mobile screen constraints and online browsing habits (Ryan, 2018), these ranking lists are presented vertically. However, for product merchants, presenting category rankings does not require including all items in that category, and product or brand rankings more commonly adopt TOP10 formats, such as JD.com' s gaming phone rankings or phone performance leaderboards, which are typically displayed horizontally to consumers. This indicates that for ranking lists with a moderate number of items, both merchants and third-party organizations can switch between vertical and horizontal presentation formats. Nevertheless, how consumers evaluate items in these two different presentation formats and the underlying mechanisms remain unknown. Addressing this gap, the present research examines the theoretical and practical significance of how different ranking list presentation formats (vertical vs. horizontal) influence consumer evaluation differences.

1.1 Ranking Effects

Consumers process information from ranking lists displayed in different formats differently. Isaac and Schindler (2014) demonstrated that consumers show greater preference for improvements that cross integer boundaries (e.g., moving from 11 to 10) compared to equivalent improvements within the same tier (e.g., from 12 to 11 or 10 to 9), a phenomenon termed the "top-ten effect." Leclerc et al. (2005) found that when making product decisions, consumers rely primarily on favorable numerical rankings within a specific list without considering the overall standing of the ranking list itself. Sevilla et al. (2018) examined differences between merchants using numerical rankings (e.g., "top 10" out of 50 products) versus percentage rankings (e.g., "top 20%" out of 50 products), finding that consumers respond more positively to numerical rankings when the list contains fewer items (<100). In summary, previous research on rankings has focused on how differences in numerical presentation formats influence consumer evaluations, with limited investigation into ranking lists containing the same number of items but displayed in different formats.

1.2 Number Spatial Position Effect

Previous research on how individuals cognitively process numerical information in different spatial positions has primarily focused on the relationship between spatial position and numerical magnitude estimation. For instance, individuals perceive numbers on the right as larger and numbers on the left as smaller (Dehaene et al., 1999), because people subconsciously process numbers using a mental number line (Dehaene et al., 2003), counting in ascending order from left to right, which influences numerical estimation (Cai et al., 2012). Additionally, numbers in vertical positions also affect magnitude perception. Hartmann et al. (2014) investigated top-down spatial-numerical associations, demonstrating that top positions are associated with larger numbers while bottom positions are associated with smaller numbers (Hartmann et al., 2014). Unlike previous studies that focused solely on the link between spatial position and numerical magnitude estimation, the present research aims to explore how different positions (horizontal vs. vertical) in ranking list contexts influence numerical processing mechanisms and how numerical and spatial position elements jointly affect consumer evaluation processes.

1.3 Polarization Effect

Attitude polarization refers to the phenomenon where prior attitudes become more extreme after exposure to attitude-consistent information (Abelson, 1995). Most research based on attitude polarization theory has investigated the causes of evaluation polarization. Previous studies have attributed consumer evaluation polarization primarily to emotional and cognitive factors. From an emotional perspective, individuals' arousal emotions affect evaluation polarization. For example, Zhu et al. (2015) demonstrated that scarcity-induced arousal leads to more polarized product evaluations and influences consumers' subsequent diversification choices. From a cognitive perspective, research has focused on how attention drives evaluation polarization. Rodas and Roedder (2019) showed that concentrated attention makes thoughts about a product repeatedly accessible, thereby making evaluations of liked products more positive. In summary, the psychological drivers of evaluation polarization have been examined from both cognitive and emotional perspectives. The present research primarily focuses on how spatial contextual cues in ranking list presentation formats drive consumers' polarized evaluations of list items.

1.4 The Influence of Ranking List Presentation Formats on Item Evaluation Polarization

Spatial metaphors encompass two dimensions: static spatial position metaphors and dynamic spatial distance metaphors (Liu & Zhang, 2019). Based on static spatial position metaphors, we argue that vertical ranking lists increase perceived differences among items compared to horizontal ranking lists for two reasons. First, in metaphorical judgment research, verticality is associated with dominance. Most studies indicate that dominance is linked to higher positions in vertical space, while submissiveness is linked to lower positions (Lakoff & Johnson, 1999; Schwartz, 1981). This dominance association enables individuals to connect abstract concepts with vertical experiences. For example, verticality helps individuals construct abstract concepts such as "promotion" and "demotion," as reflected in metaphors like "good is up" and "bad is down" (e.g., Bergen et al., 2007; Casasanto, 2009; Meier et al., 2007; Wu et al., 2006). Second, regarding static spatial positions, the hierarchy of spatial height in visual fields is: vertical dimension "up (high)" > horizontal dimension "left and right" > vertical dimension "down (low)." According to spatial position metaphors of relative height, evaluations of items in vertical "up (high)" positions are superior to those in horizontal "left" positions, while evaluations of items in horizontal "right" positions are superior to those in vertical "down (low)" positions. Thus, from the perspective of static spatial metaphors, perceived differences among items in vertical ranking lists are greater than those in horizontal ranking lists.

From the perspective of dynamic spatial distance metaphors, we also argue that vertical ranking lists increase perceived differences among items compared to horizontal ranking lists. First, Williams (1966) demonstrated that people process horizontally arranged numerical information faster than vertical numerical information because the muscles controlling horizontal eye movement are stronger than those controlling vertical movement (Cogan, 1949), making horizontal information processing easier and more fluent (Deng et al., 2016). This processing ease and fluency result in shorter processing times for the same number of items. According to the mapping relationship between spatial distance and time length (Landau et al., 2010), shorter processing times for horizontally displayed items lead to shorter perceived spatial distances between items. Second, according to the elongation hypothesis (Holmberg, 1975), individuals assign greater perceptual salience to vertical dimensions, thus weighting vertical dimensional information more heavily in decision-making and perceiving vertical length as longer than horizontal length (Brosvic & Cohen, 1988). This leads individuals to perceive greater spatial distances between vertically arranged numbers than horizontally arranged numbers. Based on Chae et al. (2013), merely altering the spatial proximity between product images in advertisements influences judgments of product similarity because spatial proximity is associated with causality (Faro, 2010). This resembles the processing mechanism in categorization, where individuals place similar category items in adjacent spaces, meaning spatial proximity leads to similarity inferences (Bargh et al., 2010). Therefore, when individuals perceive larger spatial distances between items in vertical lists, they infer lower similarity between the ranked items, resulting in greater perceived differences between the two items.

In summary, based on both spatial position and spatial distance metaphors, we infer that individuals perceive greater differences among items in vertical ranking lists than in horizontal ranking lists, and these differences lead to greater evaluation differences when evaluating corresponding items. We term this the "evaluation polarization effect." Accordingly, we propose the following hypotheses:

H1: Compared to horizontally displayed ranking lists, consumers exhibit greater evaluation polarization for items in vertical ranking lists.

H2: Perceived item differences mediate the relationship between ranking list presentation format and evaluation polarization.

1.5 The Moderating Role of Attribute Evaluability

Evaluability refers to the ease and confidence with which individuals can make subjective judgments about the superiority or inferiority of an objective attribute (Hsee, 1996; Hsee & Zhang, 2010). General evaluability theory proposes three important influencing factors from a reference system perspective: (1) nature, which is intrinsic reference information reflecting long-evolved human reference systems; (2) knowledge, which is learned reference information based on whether individuals have accumulated value distributions of target attributes as reference information; and (3) mode, which is immediate reference information depending on whether other options serve as references. Previous research indicates that if any one factor has high evaluability, the overall target has high evaluability (Lu & Hsee, 2018). Attribute evaluability produces ranking effects (Hsee, 1998) because evaluable attributes provide reference rules for ranking. When consumers filter choice sets, they typically sort by product attributes (e.g., comprehensive score, product performance, sales volume). When ranking lists do not provide specific attribute scores for ranked products, consumers only know the priority order and cannot accurately estimate the actual value gaps between products, which increases the difficulty of judging specific differences between items and leads to evaluation bias. Conversely, when attribute evaluability is high, individuals can accurately calculate the actual gaps between each rank, effectively eliminating polarization effects caused by evaluation bias. Accordingly, we propose:

H3: Attribute evaluability moderates the mediating effect of ranking list presentation format on evaluation polarization. The evaluation polarization effect emerges only when the evaluability of attributes displayed in the ranking list is low; when attribute evaluability is high, the evaluation polarization effect disappears.

Based on hypotheses H1–H3, we constructed the hypothesized model shown in Figure 1 [Figure 1: see original paper]. To test these three hypotheses, we conducted five experiments, including two online experiments and three laboratory experiments. The specific experimental logic and details are presented in Table 1 .

Experiment 1a

Purpose. Experiment 1a aimed to preliminarily examine how presentation format (horizontal vs. vertical) of positive ranking lists influences evaluation differences among ranked items.

Method. Experiment 1a employed an online questionnaire administered through the Credamo platform, recruiting 200 participants. The experiment used a mixed design of 2 (ranking list presentation format: horizontal vs. vertical, between-subjects) × 2 (item rank: 1st vs. 10th, repeated measures). Participants were randomly assigned to either the horizontal or vertical ranking list group. They were informed that ZOL.com, as a professional third-party ranking institution, had released the top 10 mobile phones by UI performance for the first half of 2019 (horizontal vs. vertical, see Appendix A). After viewing the ranking list for 30 seconds , participants answered questions. First, they estimated the comprehensive evaluation scores (0–100) for the 1st and 10th ranked phones. Second, they reported their familiarity with these brands (Zhu, 2019). Finally, they completed demographic questions and received compensation.

Participants. The experiment recruited 200 participants, including 115 males (57.50%) and 85 females (42.50%), with a mean age of 28.45 years (SD = 7.92). Using G*Power 3.1 for one-way ANOVA (Faul et al., 2009), with two groups, df = 1, effect size (f) = 0.4, and significance level = 0.05, the power exceeded 0.99, surpassing the baseline of 0.80, indicating adequate statistical power.

Results. A repeated-measures ANCOVA with ranking list presentation format as the between-subjects factor, item rank as the within-subjects factor, and brand familiarity and device type as covariates revealed significant main effects of presentation format (F(1, 194) = 6.88, p = 0.021, ^{2} = 0.08) and item rank (F(1, 194) = 20.04, p < 0.001, ^{2} = 0.10), and a significant interaction effect (F(1, 194) = 9.94, p < 0.001, ^{2} = 0.06). To further verify the effect of presentation format on evaluation differences, we conducted two analyses. First, we calculated the difference between estimated scores for the 1st and 10th ranks and compared these differences across formats. The evaluation difference for vertical lists (M_{vertical}-difference = 47.41, SD = 15.66) was significantly greater than for horizontal lists (M_{horizontal}-difference = 24.04, SD = 14.02; F(1, 194) = 18.77, p < 0.001, ^{2} = 0.04), as shown in Figure 2 [Figure 2: see original paper].

Second, to identify the source of this evaluation difference, we compared scores for the 1st and 10th ranks separately across formats. When estimating scores for the 1st-ranked phone, participants gave significantly higher ratings to the vertical list (M_{vertical}-1st = 94.42, SD = 13.22) than to the horizontal list (M_{horizontal}-1st = 80.73, SD = 10.93; F(1, 194) = 9.17, p = 0.009, ^{2} = 0.04) (see Figure 3 [Figure 3: see original paper]a). Conversely, when estimating scores for the 10th-ranked phone, participants gave significantly lower ratings to the vertical list (M_{vertical}-10th = 47.01, SD = 9.03) than to the horizontal list (M_{horizontal}-10th = 56.69, SD = 12.12; F(1, 194) = 13.07, p < 0.001, ^{2} = 0.08), as shown in Figure 3b. Thus, the evaluation polarization effect is jointly driven by differences in both the top and bottom ranked items. Notably, the contribution of the 10th-ranked item to this effect was more significant than that of the 1st-ranked item, likely due to a ceiling effect for top-ranked items in positive ranking lists. These results provide preliminary evidence for H1 in the context of positive ranking lists.

Experiment 1b

Purpose. Experiment 1b aimed to preliminarily examine how presentation format (horizontal vs. vertical) of negative ranking lists influences evaluation differences among ranked items. To eliminate stimulus interference, we used the same stimuli as Experiment 1a but with different descriptions. Additionally, because Experiment 1a used an online questionnaire where participants responded on either mobile phones or PCs, and although we controlled for stimulus aspect ratios and clarity and included device type as a covariate, screen size differences between mobile and PC could not be fundamentally eliminated. Therefore, in Experiment 1b, we included device type (PC vs. mobile) as a between-subjects factor.

Method. Experiment 1b employed an online questionnaire administered through a marketing research platform, recruiting 360 participants. The experiment used a mixed design of 2 (presentation format: horizontal vs. vertical, between-subjects) × 2 (device type: PC vs. mobile, between-subjects) × 2 (item rank: 10th from bottom vs. 1st from bottom, repeated measures). Regarding device type manipulation, upon entering the questionnaire, participants saw a prompt: "This questionnaire can only be completed on a PC (or mobile device). Responses from incorrect devices will be considered invalid. Device type will be verified after completion, and invalid responses will not receive compensation." The presentation format manipulation was identical to Experiment 1a, except participants were told that ZOL.com, as a professional third-party ranking institution, had released the bottom 10 phones by UI performance for the first half of 2019 to help consumers avoid poor-performing phones. The vertical group saw the list from top to bottom as 10th-from-bottom to bottom-ranked brands, while the horizontal group saw it from left to right. After viewing the list for 30 seconds, participants first confirmed their device type, then estimated comprehensive scores (0–100) for the 10th-from-bottom and bottom-ranked brands, reported brand familiarity (Zhu, 2019), completed demographics, and received compensation.

Participants. After screening for device consistency between assigned condition and actual device used[1], we obtained 332 valid responses (149 males, 44.88%; 183 females, 55.12%; mean age = 25.37 years, SD = 6.11). G*Power 3.1 analysis indicated adequate statistical power.

Results. A repeated-measures ANCOVA with presentation format and device type as between-subjects factors, item rank as within-subjects factor, and brand familiarity as covariate revealed significant main effects of presentation format ($F(1, 323) = 47.05$, $p < 0.001$, $^2 = 0.08$) and item rank ($F(1, 323) = 26.98$, $p < 0.001$, $^2 = 0.06$), but no significant main effect of device type ($F(1, 323) = 2.05$, $p = 0.171$). The interaction between presentation format and item rank was significant (F(1, 323) = 13.28, p < 0.001, ^{2} = 0.04), while other interactions were non-significant (format × device: F(1, 323) = 3.19, p = 0.124; rank × device: F(1, 323) = 1.88, p = 0.205; three-way interaction: F(1, 323) = 2.995, p = 0.149). These results indicate that evaluations were not affected by PC or mobile device type.

Following the same analytical approach as Experiment 1a, we first compared evaluation differences across formats. The difference between 10th-from-bottom and bottom-ranked items was significantly greater for vertical lists (M_{vertical}-difference = 21.19, SD = 5.92) than horizontal lists (M_{horizontal}-difference = 4.56, SD = 4.11; F(1, 323) = 68.29, p < 0.001, ^{2} = 0.10), as shown in Figure 4 [Figure 4: see original paper].

Second, we compared scores separately. For the 10th-from-bottom rank, vertical list ratings (M_{vertical}-10th = 32.17, SD = 9.88) were significantly higher than horizontal list ratings (M_{horizontal}-10th = 21.08, SD = 8.04; F(1, 323) = 11.37, p = 0.011, ^{2} = 0.06) (see Figure 5 [Figure 5: see original paper]a). Conversely, for the bottom-ranked item, vertical list ratings (M_{vertical}-bottom = 10.98, SD = 7.66) were marginally lower than horizontal list ratings (M_{horizontal}-bottom = 16.52, SD = 8.54; F(1, 323) = 7.03, p = 0.082, ^{2} = 0.02), as shown in Figure 5b. Thus, the evaluation difference is jointly driven by both the 10th-from-bottom and bottom-ranked items, with the 10th-from-bottom item contributing more significantly. The marginal difference for the bottom-ranked item likely reflects a floor effect in negative ranking lists. These results replicate Experiment 1a, showing that vertical presentation produces greater evaluation polarization between the 10th-from-bottom and bottom-ranked items, providing preliminary evidence for H1 in negative ranking list contexts.

Experiment 2

Purpose. Experiment 2 used a laboratory experiment with a gift product ranking scenario to expand external validity while controlling item values. To examine evaluation difference trends across lists with varying numbers of items, participants evaluated all items in the ranking list. We used ranking lists containing 10 items.

Method. Two hundred forty participants were recruited for a 2 (gift ranking presentation: horizontal vs. vertical, between-subjects) × 10 (target gift rank: 1–10, repeated measures) mixed design. Participants were told there were only 10 gifts of equal value, ranked by market research from most popular (1) to least popular (10). They were randomly assigned to vertical or horizontal ranking list groups. The questionnaire had two parts: (1) dependent variable measurement—participants estimated scores (1–100) for each gift in strict rank order to avoid order effects; (2) demographics and suspicion check, followed by compensation.

Participants. After excluding participants who guessed the purpose, 228 valid responses remained (horizontal: n = 109; vertical: n = 119; 92 males, 40.35%;

136 females, 59.65%; mean age = 22.18 years, SD = 3.07). G*Power 3.1 indicated adequate statistical power.

Results. Analysis proceeded in three steps:

Step 1: Mixed Linear Model. We constructed a mixed linear model with presentation format as between-subjects factor, the 10 gifts as continuous within-subjects factor, and their interaction predicting evaluation trends. The model included random intercepts and slopes for the number of gifts to account for repeated measures. Results showed evaluation scores decreased as rank increased (F(1, 226) = 105.27, p < 0.001). The main effect of presentation format was non-significant (F(1, 226) = 2.57, p = 0.147), but the format × rank interaction was significant (F(1, 226) = 7.92, p = 0.014, ^{2} = 0.08). Further analysis revealed that in horizontal lists, evaluation scores declined more slowly ($\beta = -2.44$) than in vertical lists ($\beta = -3.89$; t(227) = 4.98, p = 0.028). Thus, as the number of evaluated items increased, the difference between top and bottom item evaluations was significantly greater in vertical than horizontal lists, suggesting the polarization effect strengthens with more items.

Step 2: Repeated Measures ANOVA. Given that marketers commonly use 5- or 10-item rankings, we focused on these list lengths. Key findings were:

First, we examined evaluation differences at specific ranks (1, 5, and 10) across formats. Vertical list participants rated the 1st item (M_{vertical}-1st = 93.05, SD = 4.29) significantly higher than horizontal list participants (M_{horizontal}-1st = 86.90, SD = 5.69; F(1, 323) = 18.03, p = 0.021, ^{2} = 0.07). Conversely, vertical list participants rated the 10th item (M_{vertical}-10th = 11.39, SD = 6.31) significantly lower than horizontal list participants (M_{horizontal}-10th = 22.44, SD = 5.99; F(1, 323) = 38.92, p < 0.001, ^{2} = 0.10). No difference emerged for the 5th item (M_{vertical}-5th = 56.69, SD = 11.10 vs. M_{horizontal}-5th = 55.73, SD = 12.92; F(1, 323) = 0.87, p = 0.891). This validates the static spatial position metaphor mechanism: the middle item (rank 5) occupies equivalent horizontal visual positions across formats, yielding no evaluation difference, while vertical top positions are evaluated more favorably and vertical bottom positions more negatively than their horizontal counterparts.

Second, we compared top-bottom evaluation differences across formats. The difference between 1st and 10th items was significantly greater in vertical lists (M_{vertical}-top-bottom = 81.67, SD = 15.62) than horizontal lists (M_{horizontal}-top-bottom = 64.46, SD = 13.75; F(1, 323) = 52.49, p < 0.001, ^{2} = 0.14), supporting the dynamic spatial distance metaphor mechanism.

Third, we examined how list length affects evaluation differences. For 5-item lists, the top-bottom difference was marginally greater in vertical lists (M_{vertical}-5item = 36.37, SD = 6.04) than horizontal lists (M_{horizontal}-5item = 31.17, SD = 5.88; F(1, 323) = 3.89, p = 0.055, ^{2} = 0.04). For 10-item lists, this difference became more pronounced (vertical: M = 81.67, SD = 15.62; horizontal: M = 64.46, SD = 13.75; F(1, 323) = 52.49, p < 0.001, ^{2} = 0.14).

However, with very short lists (2 items: $F(1, 323) = 1.07$, $p = 0.694$; 3 items: $F(1, 323) = 1.83$, $p = 0.311$), the polarization effect disappeared.

Step 3: Mediation Analysis. We tested whether mean differences between adjacent items mediated the effect of presentation format on cumulative top-bottom evaluation differences. Using Hayes' (2013) bootstrap method with Process Model 4 (10,000 samples), we found significant mediation (LLCI = 0.063, ULCI = 0.189, excluding 0) with an indirect effect of 0.121, providing preliminary evidence for the mediating role of perceived item differences.

Discussion. Experiment 2 replicated the main effect across ranking list formats (full-item, positive, and negative rankings) and demonstrated that polarization strengthens with more items. The mediation analysis provided preliminary evidence for perceived item differences, though this cannot serve as primary theoretical evidence because perceived differences differ from evaluation differences and cannot exclude alternative explanations. Therefore, Experiment 3 further investigated the underlying mechanism.

Experiment 3

Purpose. Experiment 3 tested the mediating role of perceived item differences using a common school ranking scenario with 10 items.

Method. Two hundred participants were recruited in a university business school laboratory for a 2 (presentation format: vertical vs. horizontal) × 2 (item rank: 1 vs. 10, repeated measures) mixed design. Participants were randomly assigned to vertical or horizontal groups. The scenario described an international exchange program partnership with 10 business schools ranked by *Bloomberg Businessweek* (see Appendix B). The questionnaire included: (1) evaluation polarization measured through comprehensive score estimation (0–100 for 1st and 10th schools) and attitude preference (exchange willingness on a 7-point scale)[2]; (2) perceived item differences measured with two items adapted from Bao et al. (2011) ("The comprehensive strength gap between business schools is large," "Significant differences exist between business schools," $r = 0.74$, $p < 0.001$); (3) alternative explanations of emotion and arousal, measured because Crawford et al. (2006) linked verticality with affect (e.g., "high spirits," "feeling down"). Emotion was measured with three items (1 = very negative to 7 = very positive; 1 = very unhappy to 7 = very happy; 1 = very calm to 7 = very anxious; Cronbach' s $\alpha = 0.84$) (Ding & Zhong, 2020). Arousal was measured with three items ("felt energetic," "felt excited," "felt sleepy [reverse-coded]" ; Cronbach' s $\alpha = 0.78$) (Koo & Lee, 2011); (4) demographics and suspicion check.

Participants. One hundred eighty-seven valid participants remained (86 males, 45.99%; 101 females, 54.01%; mean age = 23.17 years, SD = 4.06). G*Power 3.1 indicated adequate statistical power.

Results. Perceived item differences. Participants in vertical lists perceived significantly greater differences among business schools (M_{vertical} = 5.04, SD = 1.27) than those in horizontal lists (M_{horizontal} = 4.16, SD = 1.51; F(1, 185) = 17.54, $p < 0.001$, $^{2} = 0.08$).

Evaluation polarization. For exchange willingness, vertical list participants showed greater willingness for the 1st-ranked school (M_{vertical}-1st = 5.67, SD = 1.13) than horizontal list participants (M_{horizontal}-1st = 4.81, SD = 1.32; F(1, 185) = 7.95, p = 0.013, $^{2} = 0.04$), and lower willingness for the 10th-ranked school (M_{vertical}-10th = 3.42, SD = 1.75 vs. M_{horizontal}-10th = 4.30, SD = 1.19; F(1, 185) = 11.258, $p < 0.001$, $^{2} = 0.10$). Similar patterns emerged for comprehensive scores (vertical 1st: M = 94.77, SD = 10.21 vs. horizontal 1st: M = 86.55, SD = 9.16, F(1, 185) = 8.36, p = 0.008, $^{2} = 0.06$; vertical 10th: M = 56.98, SD = 18.04 vs. horizontal 10th: M = 71.22, SD = 14.77, F(1, 185) = 16.59, $p < 0.001$, $^{2} = 0.08$), replicating H1.

Mediation analysis. Using Hayes' (2013) Process Model 4 with 5,000 bootstrap samples, perceived item differences significantly mediated the effect of presentation format on both exchange willingness difference (b = 0.202, SE = 0.082, 95% CI = [0.067, 0.305], excluding 0) and comprehensive score difference (b = 0.173, SE = 0.064, 95% CI = [0.030, 0.348], excluding 0), supporting H2 (see Figures 6 [Figure 6: see original paper] and 7 [Figure 7: see original paper]).

Alternative explanations. ANOVAs revealed no significant differences in emotion (F(1, 185) = 1.09, p = 0.342) or arousal (F(1, 185) = 1.84, p = 0.204) across formats. Although perceived differences correlated with emotion (r = 0.32, $p < 0.001$) and arousal (r = 0.49, $p < 0.001$), neither emotion (indirect effect = 0.041, SE = 0.028, 95% CI = [−0.077, 0.121]) nor arousal (indirect effect = 0.059, SE = 0.092, 95% CI = [−0.137, 0.098]) mediated the effect. Controlling for these variables, the main effect remained significant (F(3, 183) = 7.98, p = 0.015) and perceived differences still mediated the effect (exchange willingness: indirect effect = 0.182, SE = 0.104, 95% CI = [0.087, 0.285]; comprehensive score: indirect effect = 0.145, SE = 0.072, 95% CI = [0.040, 0.194]), effectively ruling out these alternatives.

Discussion. Experiment 3 replicated the main effect and directly measured perceived item differences as the mediator. However, boundary conditions remained unclear, prompting Experiment 4.

Experiment 4

Purpose. Experiment 4 examined the boundary condition of attribute evaluability. Since evaluation polarization stems from cognitive bias induced by spatial position cues, providing reference information to increase attribute evaluability should eliminate this bias. We used sales volume—a stable, objective attribute—to manipulate evaluability.

Method. Three hundred participants were recruited for a 2 (presentation format: vertical vs. horizontal, between-subjects) × 2 (attribute evaluability: high vs. low, between-subjects) × 2 (item rank: 1 vs. 10, repeated measures) mixed design. Participants were randomly assigned to four groups. The low evaluability condition used the same stimuli as Experiment 1a, informing participants the list was sorted by monthly sales volume without providing specific sales figures. The high evaluability condition added specific monthly sales numbers (e.g., 387 units for 1st rank, 304 for 2nd). After viewing stimuli, participants completed: (1) an evaluability manipulation check ("Evaluating the ranking information is not difficult," "I can quickly evaluate the information," "Evaluating requires little effort" ; Cronbach' s $\alpha = 0.88$) (Sun, 2011); (2) brand attitude ratings (7-point scale) for 1st and 10th items; (3) perceived sales differences among the 10 brands; (4) demographics.

Participants. After excluding 21 participants who failed attention checks or had duplicate IP addresses, 279 valid participants remained (129 males, 46.24%; 150 females, 53.76%; mean age = 22.08 years, SD = 4.21). G*Power 3.1 indicated adequate statistical power.

Results. Manipulation check. Attribute evaluability was significantly higher when specific scores were provided (M_{high} = 5.18, SD = 1.44) than when not provided (M_{low} = 3.04, SD = 2.02; F(1, 277) = 94.02, p < 0.001, ^{2} = 0.10), confirming successful manipulation.

Perceived item differences. Under low evaluability, vertical lists produced significantly greater perceived differences (M_{vertical} = 5.08, SD = 1.42) than horizontal lists (M_{horizontal} = 4.21, SD = 1.88; F(1, 271) = 14.27, p < 0.001, ^{2} = 0.08). Under high evaluability, no significant difference emerged (F(1, 271) = 3.09, p = 0.218).

Evaluation polarization. Under low evaluability, the evaluation difference between 1st and 10th items was significantly greater for vertical lists (M_{vertical}-difference = 2.81, SD = 0.75) than horizontal lists (M_{horizontal}-difference = 1.91, SD = 1.08; F(1, 271) = 44.37, p < 0.001, ^{2} = 0.09). Under high evaluability, this difference disappeared (F(1, 271) = 2.98, p = 0.120) (see Figure 8 [Figure 8: see original paper]).

Moderated mediation. Using Hayes' (2013) PROCESS Model 8 with 5,000 bootstrap samples, we found significant moderated mediation (b = 0.512, SE = 0.231, 95% CI = [0.145, 0.781], excluding 0). Specifically, under low evaluability, perceived differences mediated the effect (b = 0.482, SE = 0.110, 95% CI = [0.128, 0.553], excluding 0), whereas under high evaluability, this mediation was non-significant (b = 0.030, SE = 0.084, 95% CI = [−0.098, 0.125], including 0), supporting H3.

Discussion. Experiment 4 confirmed that evaluation polarization occurs only under low attribute evaluability and disappears when evaluability is high.

General Discussion

This research investigated how different presentation formats (vertical vs. horizontal) of identical ranking lists influence consumer evaluations and the underlying mechanisms through five experiments.

Theoretical Contributions

First, this research extends the theoretical framework of ranking effects by examining how spatial position changes of ranked items influence consumer evaluations, expanding ranking list effects into the spatial context dimension. Previous research focused on how numerical elements (e.g., ranking improvements, boundary effects, numerical vs. percentage formats) influence consumer decisions, neglecting spatial position elements. Our study integrates spatial and numerical elements in ranking contexts, shifting from a "number-position" association paradigm to a "number + position → cognition" causal paradigm. We demonstrate that identical numbers in different spatial positions create cognitive biases, extending numerical cognitive bias theory beyond different number forms (e.g., odd/even, ending in 0 or 9, integers vs. decimals) to the same numerical representation dimension.

Second, we expand the application scenarios of spatial metaphor theory. While horizontal and vertical positioning are widely used in retail displays, few studies have applied position effects to ranking lists. Moreover, previous spatial metaphor research typically employed only one sub-dimension—either static position metaphors (linking verticality to power and morality) or dynamic distance metaphors (focusing on distance effects). This study systematically integrates both dimensions (static position + dynamic distance) within a unified framework to explain how ranking list spatial representation influences evaluation polarization.

Third, we extend the psychological antecedents of evaluation polarization from a cognitive perspective. While previous research examined polarization through attention (cognitive) and arousal (emotional) mechanisms, we identify perceived item differences as a novel cognitive driver. By ruling out emotional and arousal alternatives, we demonstrate that spatial contextual cues in ranking lists drive polarization through cognitive processes. Additionally, whereas previous polarization research focused on attitude or rating changes for a single product, we compare evaluation differences across different target products, expanding the manifestations and applications of evaluation polarization.

Practical Implications

First, while firms may use imprecise ranking claims to obscure unfavorable positions (Sevilla & Bagchi, 2018), this may backfire (Darke & Ritchie, 2007). Our findings suggest that horizontal presentation can mitigate negative evaluations of low-ranked products. Second, third-party ranking institutions (non-profit, aiming to help consumers identify quality) can use vertical lists, while merchants promoting their own products (e.g., JD.com self-operated products) may benefit from horizontal formats. Finally, as this is a cognitive bias, consumers and policymakers can implement interventions to reduce it, such as seeking specific score values (rank order + specific scores) to increase attribute evaluability and eliminate bias.

Future Directions

This research has limitations. First, all experiments controlled for scrolling, making conclusions applicable only to moderate-length lists (e.g., $ $10 items). Future research should examine lists with more items. Second, we used static ranking list images, whereas real browsing involves scrolling. Future studies should investigate whether the effect holds when lists are scrollable. Third, while we focused on spatial position cues with identical numbers, future research could control for numerical elements to examine how marketing or social factors influence ranking list decisions.

References

Abelson, R. P. (1995). Attitude extremity. In R. E. Petty, & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 25—42). Mahwah, NJ: Lawrence Erlbaum.

Bargh, J. A., Williams, L. E., Huang, J. Y., Song, H., & Ackerman, J. M. (2010). From the physical to the psychological: Mundane experiences influence social judgment and interpersonal behavior. *Behavioral & Brain Sciences*, 33(4), 267-268.

Bagchi, R., & Derick, F. D. (2016). Numerosity and consumer decision making. *Current Opinion in Psychology*, 10(5), 89-93.

Bagchi, R., & Ince, E. C. (2016). Is a 70% forecast more accurate than a 30% forecast? How level of a forecast affects inferences about forecasts and forecasters. *Journal of Marketing Research*, 53(1), 31-45.

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, 31(5), 733-764.

Brosvic, G. M., & Cohen, B. D. (1988). The horizontal-vertical illusion and knowledge of results. *Perceptual and Motor Skills*, 67(2), 463-469.

Bao, Y. C., Bao Y. Q., & Sheng S. B (2011). Motivating purchase of private brands: Effects of store image, product signatureness, and quality variation. *Journal of Business Research*, 64(2), 220-226.

Casasanto, & Daniel. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General*, 138(3), 351-367.

Cai, F., Shen, H., Hui, & Michael, K. (2012). The effect of location on price estimation: Understanding number-location and number-order associations. *Journal of Marketing Research*, 49(5), 718-724.

Chae, B., Xiuping, L. I., & Zhu, R. (2013). Judging product effectiveness from perceived spatial proximity. *Journal of Consumer Research*, 40(2), 317-335.

Chen, S.S., Ke, Y. N., Jiang, J., Xiao X. (2014). The influence of vertical metaphor of power on power judgment. *Journal of Psychological Science*, (2), 388-393.

Cogan, D. G. (1949). Neurology of the ocular muscles. *Journal of Nervous & Mental Disease*, 109(2), 187.

Crawford, L. E., Margolies, S. M., Drake, J. T., & Murphy, M. E. (2006). Affect biases memory of location: Evidence for the spatial representation of affect. *Cognition and Emotion*, 20(8), 1153-1169.

Darke, P. R., & Ritchie, R. J. B. (2007). The defensive consumer: Advertising deception, defensive processing, and distrust. *Journal of Marketing Research*, 44(1), 114-127.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, 284(5416), 970-974.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3-6), 487-506.

Deng, X., Kahn, B. E., Unnava, H. R., & Lee, H. (2016). A "wide" variety: Effects of horizontal versus vertical display on assortment processing, perceived variety, and choice. *Journal of Marketing Research*, 53(5), 682-698.

Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology*, 87(5), 586-598.

Ding, Y., & Zhong, J. Q. (2020). The effect of social crowding on individual preference for self-improvement products. *Acta Psychologica Sinica*, 52(2), 216-228.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses.* Behavior Research Methods*, 41(4), 1149-1160.

Faro, D. (2010). Changing the future by reshaping the past: The influence of causal beliefs on estimates of time to onset. *Journal of Consumer Research*, 37(2), 279-291.

Gamliel, E., & Peer, E. (2016). The average fuel-efficiency fallacy: Overestimation of average fuel efficiency and how it can lead to biased decisions. *Journal of Behavioral Decision Making*, 30(2), 435-445.

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*.

Hartmann, M., Gashaj, V., Stahnke, A., & Mast, F. W. (2014). There is more than "more is up" : Hand and foot responses reverse the vertical association of number magnitudes. *Journal of Experimental Psychology: Human Perception & Performance*, 40(4), 1401–1414.

Holmberg, L. (1975). The influence of elongation on the perception of volume of geometrically simple objects. *Psychological Research Bulletin*, 15(2), 1–18.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3), 247–257.

Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, 11(2), 107–121.

Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, 5(4), 343–355.

Isaac, M. S., & Schindler, R. M. (2014). The top-ten effect: Consumers' subjective categorization of ranked lists. *Journal of Consumer Research*, 40(6), 1181–1202.

Isaac, M. S., Brough, A. R., & Grayson, K. (2016). Is top 10 better than top 9? The role of expectations in consumer response to imprecise rank claims. *Journal of Marketing Research*, 53(3), 338–353.

Judd, C. M., & Lusk, C. M. (1984). Knowledge structures and evaluative judgments: Effects of structural variables on judgmental extremity. *Journal of Personality and Social Psychology*, 46(6), 1193–1207.

Koo, D. M., & Lee, J. H. (2011). Inter-relationships among dominance, energetic and tense arousal, and pleasure, and differences in their impacts under online vs. offline environment. *Computers in Human Behavior*, 27(5), 1740–1750.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenges to Western thought*. New York: Basic Books.

Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, 136(6), 1045–1067.

Leclerc, F., Hsee, C. K., & Nunes, J. C. (2005). Narrow focusing: Why the relative position of a good in its category matters more than it should. *Marketing Science*, 24(2), 194–205.

Liberman, A., & Chaiken, S. (1991). Value conflict and thought-induced attitude change. *Journal of Experimental Social Psychology*, 27(3), 203–216.

Liu, H. Y, Zhang, S. X. (2019). To see the truth of space: A literature review and prospects of spatial metaphors' effects on consumption behavior. *Foreign*

Economics & Management, 41(2), 59-72.

Lu X., Hsee, Christopher K. (2018). Joint evaluation versus single evaluation: A field full of potentials. *Acta Psychologica Sinica*, 50(8), 827-839.

Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What' s "up" with god? Vertical space as a representation of the divine. *Journal of Personality & Social Psychology*, 93(5), 699-710.

Millar, Tesser, M. G., & Abraham. (1986). Thought-induced attitude change: The effects of schema structure and commitment. *Journal of Personality and Social Psychology*, 51(2), 259-269.

Ouellet, M., Santiago, J., Israeli, Z., & Gabay, S. (2010). Is the future the right time?. *Experimental Psychology*, 57(4), 308-314.

Paulhus, D. L., & Lim, D. (1994). Arousal and evaluative extremity in social judgments: A dynamic complexity model. *European Journal of Social Psychology*, 24(1), 89-99.

Pham, M. T. (1996). Cue representation and selection effects of arousal on persuasion. *Journal of Consumer Research*, 22(4), 373-387.

Pope, D. G. (2009). Reacting to rankings: Evidence from "America' s best hospitals" . *Journal of Health Economics*, 28(6), 1154-1165.

Romero M., & Biswas D. (2016). Healthy-left, unhealthy-right: Can displaying healthy items to the left (versus right) of unhealthy items nudge healthier choices?. *Journal of Consumer Research*, 43(1), 103-112.

Rodas, M. A., & Roedder, J. D. (2019). The secrecy effect: Secret consumption increases women' s product evaluations and choice. *Journal of Consumer Research*, 46(6), 1093-1109.

Ryan, K. M. (2018). Vertical video: Rupturing the aesthetic paradigm. *Visual Communication*, 17(2), 245-261.

Schwartz, B. (1981). *Vertical classification: A study in structuralism and the sociology of knowledge*. Chicago: University of Chicago Press.

Schubert, T. W. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality & Social Psychology*, 89(1), 1-21.

Shen, L. X., Hsee, C. K., Wu, Q. S., & Tsai, C. I. (2012). Overpredicting and underprofiting in pricing decisions: Evaluability in pricing decisions. *Journal of Behavioral Decision Making*, 25(5), 512-521.

Sevilla, J., Isaac, M. S., & Bagchi, R. (2018). Format neglect: How the use of numerical versus percentage rank claims influences consumer judgments. *Journal of Marketing*, 82(6), 99-112.

Song, Y. Q., Zhang, J. J. (2014). Temporal-spatial metaphor in conceptual representation: Can spatial information be activated when processing the temporal reference which implied in the changing shape of objects?. *Acta Psychologica Sinica*, 46(2), 216-226.

Sun, J. (2011). Consumer brand preference construction: The moderating role of evaluation mode. *Management Review*, 23(8), 103-111.

Tsiros, M. (2017). Convexity neglect in consumer decision making. *Journal of Marketing Behavior*, 2(4), 286-286.

Wang, Z., Lu Z. Y. (2013). The vertical spatial metaphor of moral concepts and its influence on cognition. *Acta Psychologica Sinica*, 45(5), 538-545.

Williams, C. M. (1966). Horizontal versus vertical display of numbers. *Human Factors*, 8(3), 237.

Wu, L. M., Mo, L., & Wang, R. M. (2006). The activation process of spatial representations during real-time comprehension of verbs. *Acta Psychologica Sinica*, 38(5), 663-671.

Yunhui, H., & Han, G. (2018). The minimal deviation effect: Numbers just above a categorical boundary enhance consumer desire. *Journal of Consumer Research*, 45(4), 775-791.

Zhang M., & Wang J. (2009). Psychological distance asymmetry: The spatial dimension vs. other dimensions. *Journal of Consumer Psychology*, 19(3), 497-507.

Zhu, M., & Ratner, R. K. (2015). Scarcity polarizes preferences: The impact on choice among multiple items in a product class. *Journal of Marketing Research*, 52(1), 13-26.

Zhao, X., He, X., Wei, Z., Chen, G., Chen, Q., & Huang, L. (2018). Interpersonal choice: The advantage on the left or on the right?. *International Journal of Psychology*, 53(5), 331-338.

Zhu, Y. M. (2019). Influence of consumer participation level on identification in online brand community—The moderating effect of product type and brand familiarity. *Journal of Business Economics*, (2), 51-61.

---

[1] Data screening principle: We retained participants whose actual device (recorded by the system) matched their assigned condition, not self-reported device, because some participants assigned to the PC condition may have used mobile devices while incorrectly reporting PC usage.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.