Cognitive Anchoring of Color Cues on Online Review Ratings

Emergent Research Forum Paper

Dario Bonaretti
Louisiana State University
dbonar1@lsu.edu

Marcin Bartosiak
University of Pavia
ml.bartosiak@universitadipavia.it

Gabriele Piccoli
Louisiana State University
gpiccoli@cct.lsu.edu

Abstract

Online review systems (ORS) such as TripAdvisor or Yelp collect numeric evaluations from reviewers using interval scales. However, the UI of interval scales differ remarkably across ORS, even though prior research suggests that design cues of the interval scale can bias individual's interpretation of the scale and thus the numeric evaluations. The impact of the UI on numeric evaluations is particularly relevant in the ORS domain since there is a tight correlation between reviews' ratings and profits. In this research-in-progress, we outline the theoretical foundation for investigating the cognitive impact of color cues in the interpretation of interval scales, and the possible distortions of numeric evaluations they could lead to.

Keywords

Online Review, Color Effect, Cognitive Anchor, Cognitive Cue, Framing Effect.

Introduction

Online review systems (ORS) are web applications (e.g., TripAdvisor, Yelp) for collecting and presenting online reviews: A structured form of user generated content that consists of both numeric and textual peer evaluations of products or services. Building on prior research on cognitive biases and anchoring effect (Tversky and Kahneman 1974), scholars investigated the effect of cognitive biases such as social anchors on user's review rating (Arazy et al. 2015). Moreover, prior research suggests that also design elements might affect how users interpret the rating scale, and eventually the numeric evaluation (De Langhe et al. 2011). The impact of rating scales' design on numeric evaluations received limited attention, even if interval scales' design varies remarkably by the number of intervals, color-design, shape, and labels. To fill this gap, our research looks at the cognitive impact of color cues, testing empirically whether color design elicits anchoring effects and eventually influence numeric evaluations on ORS.

From a practical standpoint, cognitive anchoring due to color cues could negatively impact two users' groups of ORS: (i) review readers, who increasingly ground their purchase decisions on online reviews (Chen and Xie 2008) and (ii) companies, that might use reviews to measure service quality (Duan et al. 2013). Thus, our study contributes to the field of HCI in two ways. First, by looking at how systematic biases can occur in ORS due to the design of the rating scale. In fact, recognizing whether color design distorted numeric ratings constitutes an important proof of concept to encourage further investigation in design-related biases. Second, we extend prior findings on the moderating effect of *emotional stability* on emotional anchoring due to color design.

Theoretical Framework

Anchor effect in ORS

Prior research conceptualizes anchoring as a phenomenon where "[…] different starting points yield different estimates, which are biased toward the initial value" (Tversky and Kahneman 1974, p. 1128).

Anchoring biases of users' numeric evaluations pervade ORS as well since cognitive anchors affect users' evaluation by framing the interpretation of the scale (Upshaw 1965). For example, reviewers anchor their evaluation to prior reviews (Adomavicius et al. 2011) and especially to the cumulative average rating, which expresses the general sentiment of the community and thus constitutes a *social anchor* (Arazy et al. 2015). Moreover, design elements of the scale can trigger anchor effects too. In fact, there is a systematic tendency for "more verbal intense labels on the anchoring points of a rating scale [to] lead respondents to move away from the ends of the scale" (De Langhe et al. 2011, p.367). Thus, scholars describe as *anchor contraction effect* the respondents' tendency to pick milder ratings when facing anchor cues that convey higher emotional intensity (De Langhe et al. 2011). However, personality traits such as *emotional stability* can moderate the impact of cognitive anchors (Arazy et al. 2015). Emotional stability is the "[tendency] to be more secure and self-assured" (Arazy et al. 2015, p. 49), making more emotionally stable individuals less vulnerable to the manipulative effect of anchors.

Cognitive impact of color cues

In the previous section, we presented how anchoring can bias online review rating. In this section, we articulate how color cues can trigger anchoring. We define *cue* as a perceivable and interpretable materialization of a signified. A signified is a mental concept associated to cues, such as "stop" in the presence of color red. However, the interpretation of color cues can differ depending on contextual elements. In fact, a red light is not necessary a signifier for "stop": It might mean "recording" when flickering on a camera, or "busy" when it goes on the door of an aircraft lavatory. However, it becomes arguably a signifier for "stop" when is used in combination with yellow and red in traffic lights.

Therefore, the cognitive impact of color cues depends on the interpretation of what the color cue "stands for", namely of the *signaling significance* (Jacobs 2013, p. 170) of cues about specific information (Maier et al. 2008). In spite of cross-cultural differences (Proctor and Vu 2010), individuals share a similar interpretation of the signaling value of color (Caivano 1998). For example, the dichotomy green versus red is consistent across different color models (Fehrman and Fehrman 2000), as in the case of traffic lights, which are also consistent across cultures (i.e. green="go" and red="stop"). Our review on the emotional value of colors shows that red has an arousing yet negative connotation when paired with green (Akers et al. 2012; De Bock et al. 2013; Elliot et al. 2007; Madden et al. 2000; Maier et al. 2008).

However, prior findings on color effect do not necessarily apply to color cues in UI design. In fact, we faced a gap in the existing literature between general literature on color effect, and the effect of color cues in UI design of interval scales specifically. From prior investigations on interval scale's design, we know that color cues such as color brightness can successfully manipulate individuals' numeric evaluations because respondents interpret different levels of color brightness as conveying different levels of emotional intensity (De Langhe et al. 2011). Building on these findings, we inferred that –similarly to color brightness– color hue could also portray emotional intensity when used as a cognitive cue. Thus, we circled back and mapped colors to different levels of emotional intensity referring to our review of color valence (Table 1). As from prior research, higher levels of emotional intensity should trigger an anchor contraction effect, leading respondents to select milder ratings (De Langhe et al. 2011).

The impact of online review rating

The previous section presented how color can trigger anchoring. Here we discuss why anchor effects – and distortions of review ratings broadly speaking – can hinder user's decision process. Online reviews affect sales (Park et al. 2007), and for business, it is critical to minimize negative reviews since they affect sales more heavily than positive reviews (Chevalier and Mayzlin 2006).

Moreover, individuals pay higher attention to numeric evaluations because they are a form of a peripheral cue (Fiske and Taylor 1984), and therefore require less cognitive effort during message elaboration (Petty and Cacioppo 1986). Thus review ratings constitute critical information for the decision maker to quickly discriminate between competing products or services (Filieri and McLeay 2014). For this reason, we believe that investigating the manipulative power of color design is critical to ensure that ORS constitute a reliable measurement system for collecting customer reviews.

Hypotheses Development

In the literature review, we articulated how UI design cues could affect user's behavior through anchoring. Specifically, we speculate that color cues adjust users' interpretation of the interval scale. Moreover, the magnitude of the adjustment is moderated by the emotional stability of the subject.

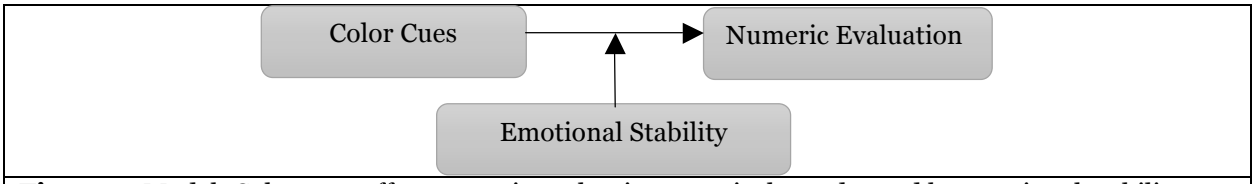

Figure 1. Model: Color cues affect numeric evaluation negatively moderated by emotional stability.

However, we discussed that the signaling significance of color cues is contextual. Thus, the magnitude and direction of the adjustment does not only depend on color effect per se but on color design as a cognitive enhancer (or reducer) of the numeric values on the interval scale. For example, green cues enhance positive emotional valence on the upper bound of the interval scale, but the effect is reversed when color green is tight to the lower bound of the scale. Thus, we hypothesize that color design presenting color with a positive valence on the upper bound, and vice versa on the lower bound of the scale, are likely to compress the distribution of the ratings towards mild ratings. However, when a monochromatic color design is used, no signaling effect occurs.

The first step to support our proposed model is showing that users adjust their interpretation of the scale based on color cues. If this holds true, we expect score adjustments to occur when color cues interact with the users' cognitive interpretation of the scale. Thus, we expect that:

H1a. Color cues displayed on-click increase the rate of change per user on the interval scale

At the same time, if color cues are given rather than displayed on-click, the adjustment should occur prior to interacting with the scale at all. Consequentially, we hypothesize that the average score will adjust but without increasing the number of changes on the interval scale, thus:

H1b. Color cues statically displayed present the same rate of change than the baseline treatment (gray)

Prior research suggests that more emotionally intense cues should elicit a central tendency effect. In particular, colors that emotionally reinforce the negativity/positivity of the ratings (i.e. red for negative ratings, and green for positives), should lead respondents to adjust towards less extreme ratings, namely showing a contraction effect.

H2a: Negative (red) color cues associated with an extremely negative rating (i.e., 1) increases the average rating of extremely negative experiences.
H2b: Positive (green) color cues associated with an extremely positive rating (i.e., 5) decreases the average rating of extremely positive experiences.

This effect due to their signaling significance does not occur when the color design does not express a meaningful color combination, such as in monochromatic design.

H2c: Monochromatic designs do not differ in the average rating between each other, but they differ compared to multi-chromatic design.

Finally, we argue that *emotional stability* moderates the effect of color cues, thus we hypothesize:

H3a: Higher levels of emotional stability reduce the increase of average ratings for extremely negative experiences.
H3b: Higher levels of emotional stability reduce the decrease of average ratings for extremely positive experiences.

Experimental Design

We test the hypotheses with a randomized six-groups posttest-only laboratory experiment. The protocol requires users to evaluate a lodging experience presented as a textual online review through a custom-

built web application (https://dds.cct.lsu.edu/ddslab/websiteAMCIS2017/). Each user evaluates from 1 (poor) to 5 (excellent) six attributes of the fictional lodging experience that would be relevant in the hospitality industry (overall, cleanliness, sleep quality, rooms, service and value). To avoid missing data, we enforce one response for each of the six attributes, and to control for changes between ratings we track user's clicks with their corresponding timestamp. The baseline treatment is *gray*, which measure the ratings in the absence of any color effect, consistently with similar studies on color perception (Elliot et al. 2007; Maier et al. 2008). *Red* and *Green* instead, test the absence of color effect when the color design is monochromatic (H2c). *Color* and *colorStatic* test the anchor contraction effect due to color effect (H2a, H2b); moreover, if *color* presents on average more changes for the same attribute compared to *colorStatic*, this would corroborate the claim that the reinterpretation actually occurs depending on the color that pops up *after* selecting the rating.

Our choice of color saturation and brightness for *red* and *green* is pragmatic, and mirrors existing real world design to ensure practical generalizability: *green* for TripAdvisor and *red* for Yelp. Note that the color from 1 in *color* is the same of 1 in *red,* and color for 5 in *color* is the same as 5 in *green*. To pursue wider generalizability on the hypothesized effect, we also include another color gradient design from Trivago (*color2*).

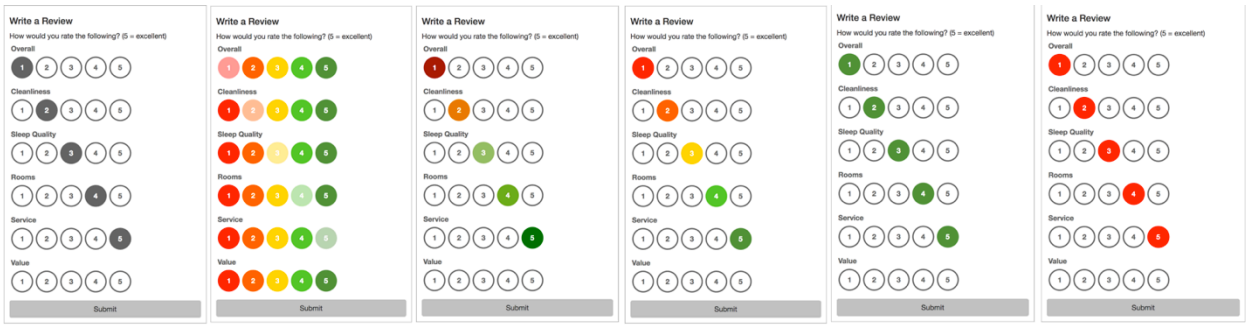

Figure 2 The six treatments, from the left to the right: *gray, colorStatic, color2, color, green* and *red.*

Finally, subjects are classified in high or low emotional stability using median split on the measure of emotional stability adapting a two items and 7-point Likert scale (Gosling et al. 2003): "I see myself as calm, emotionally stable", "I see myself as anxious, easily upset".

Discussion and future developments

This research in progress presents our preliminary understanding of the impact of UI design on numeric evaluation in the domain of ORS. Influencing numeric evaluations in ORS using color cues might be only one aspect of a more general theory of manipulation that uses design elements as cognitive anchors. Moreover, possible applications of an HCI kernel theory of manipulation through UI design are extremely relevant to the industry. For instance, anchor contraction might provide a tool for review manipulation, enabling hotels that offer poor service to inflate their reputation by reinforcing their presence on ORS that present anchor contractive UI design, or by designing their proprietary review system accordingly.

References

Adomavicius, G., Bockstedt, J., Curley, S., and Zhang, J. 2011. "Recommender systems, consumer preferences, and anchoring effects," in *RecSys 2011 Workshop on Human Decision Making in Recommender Systems*, pp. 35–42.

Akers, A., Barton, J., Cossey, R., Gainsford, P., Griffin, M., and Micklewright, D. 2012. "Visual color perception in green exercise: positive effects on mood and perceived exertion," *Environmental science & technology*, (46:16), pp. 8661–8666.

Arazy, O., Nov, O., and Kumar, N. 2015. "Personalityzation: UI personalization, theoretical grounding in HCI and design research," *AIS Transactions on Human-Computer Interaction*, (7:2), pp. 43–69.

Caivano, J. L. 1998. "Color and semiotics: A two-way street," *Color Research & Application*, (23:6), pp. 390–401.

Chen, Y., and Xie, J. 2008. "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science*, (54:3), pp. 477–491 (doi: 10.1287/mnsc.1070.0810).

Chevalier, J. A., and Mayzlin, D. 2006. "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, (43), pp. 345–354.

Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., and Riedl, J. 2003. "Is seeing believing?: how recommender system interfaces affect users' opinions," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, pp. 585–592 (available at http://dl.acm.org/citation.cfm?id=642713).

De Bock, T., Pandelaere, M., and Van Kenhove, P. 2013. "When colors backfire: The impact of color cues on moral judgment," *Journal of Consumer Psychology*, (23:3), pp. 341–348 (doi: 10.1016/j.jcps.2012.09.003).

De Langhe, B., Puntoni, S., Fernandes, D., and Van Osselaer, S. M. 2011. "The anchor contraction effect in international marketing research," *Journal of Marketing Research*, (48:2), pp. 366–380.

Duan, W., Cao, Q., Yu, Y., and Levy, S. 2013. "Mining online user-generated content: using sentiment analysis technique to study hotel service quality," in *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, IEEE, pp. 3119–3128.

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., and Meinhardt, J. 2007. "Color and psychological functioning: The effect of red on performance attainment.," *Journal of Experimental Psychology: General*, (136:1), pp. 154–168 (doi: 10.1037/0096-3445.136.1.154).

Fehrman, K., and Fehrman, C. 2000. *Color: The secret influence*, Prentice Hall.

Filieri, R., and McLeay, F. 2014. "E-WOM and accommodation an analysis of the factors that influence travelers' adoption of information from online reviews," *Journal of Travel Research*, (53:1), pp. 44–57.

Fiske, S. T., and Taylor, S. E. 1984. *Social Cognition*, New York, NY: Random House.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B. 2003. "A very brief measure of the Big-Five personality domains," *Journal of Research in personality*, (37:6), pp. 504–528.

Hill, R. A., and Barton, R. A. 2005. "Psychology: Red enhances human performance in contests," *Nature*, (435:7040), pp. 293–293.

Jacobs, G. 2013. *Comparative color vision*, Elsevier.

Madden, T. J., Hewett, K., and Roth, M. S. 2000. "Managing images in different cultures: A cross-national study of color meanings and preferences," *Journal of international marketing*, (8:4), pp. 90–107.

Maier, M. A., Elliot, A. J., and Lichtenfeld, S. 2008. "Mediation of the negative effect of red on intellectual performance," *Personality and Social Psychology Bulletin*.

Park, D.-H., Lee, J., and Han, I. 2007. "The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement," *International Journal of Electronic Commerce*, (11:4), pp. 125–148.

Petty, R. E., and Cacioppo, J. T. 1986. *The elaboration likelihood model of persuasion*, Springer.

Proctor, R. W., and Vu, K.-P. L. 2010. "Universal and culture-specific effects of display-control compatibility," *The American Journal of Psychology*, (123:4), pp. 425–435 (doi: 10.5406/amerjpsyc.123.4.0425).

Tversky, A., and Kahneman, D. 1974. "Judgment under uncertainty: Heuristics and biases," *Science*, (185:4157), pp. 1124–1131.

Upshaw, H. S. 1965. "The effect of variable perspectives on judgments of opinion statements for Thurstone scales: Equal-appearing intervals.," *Journal of Personality and Social Psychology*, (2:1), p. 60.