Context effects on beauty ratings of photos:

Building contrast effects that erode but cannot be knocked down

Cody Tousignant and Glen E. Bodner
University of Calgary

BRIEF REPORT

Correspondence to: Cody Tousignant
Department of Psychology
University of Calgary
2500 University Drive NW
Calgary AB T2N 1N4
CANADA

phone: 403-220-6682
fax: 403-282-8249
e-mail: catousig@ucalgary.ca
courier: Administration Building, Room A275
Abstract

The aesthetic value of photos is often evaluated alongside other photos (art galleries, photo albums). We examined the direction, duration, and resilience of the influence of a set of context photos on beauty ratings for a subsequent set of critical photos. Beauty ratings for a set of average-beauty photos of buildings were higher after viewing a set of low-beauty (vs. high-beauty) photos, regardless of whether the context photos depicted the same, similar, or different thematic content. This contrast effect persisted when participants were asked to avoid being influenced by the context photos, when the context photos were processed in a non-aesthetic task, and when the context and critical photos were made to seem more similar. Although we could not “knock down” these contrast effects, they “eroded” across the critical block as the groups’ recent experiences became more similar. Implications of our findings for theories of contrast/assimilation and aesthetics are discussed.

Keywords: contrast and assimilation effects, context effects, beauty ratings, aesthetic evaluations, selective-accessibility model, range-frequency model
Context effects on beauty ratings of photos:

Building contrast effects that erode but cannot be knocked down

The subjective experience of beauty has become a research focus (see Cinzia & Vittorio, 2009 and Lindell & Mueller, 2011 for reviews), within which one emphasis is on how contextual factors shape aesthetic experience (e.g., Dolese, Zellner, Vasserman, & Parker, 2005; Kirk, Skov, Hulme, Christensen, & Zeki, 2009; Locher, Smith, & Smith, 2001; Parker, Bascom, Rabinovitz, & Zellner, 2008; Silvera, Josephs, & Giesler, 2002). This contextual focus is appropriate given that we often evaluate aesthetic objects displayed alongside similar objects. For example, curators gather related pieces for exhibitions, albums of photos are posted on Facebook, and songs are grouped together into compilations and playlists. Our study examined how the beauty of an initial set of context photos (low vs. high) affects beauty judgments for a subsequent set of critical average-beauty photos. Specifically, Experiment 1 manipulated the thematic similarity of the context and critical photo sets to examine whether assimilation or contrast effects occur. Experiment 2 then examined whether we could modulate the context effects obtained in Experiment 1. Both experiments also examined the duration of context effects on beauty ratings. To our knowledge, prior studies have not measured whether context effects on aesthetic evaluations are transient or lasting, which is surprising given that aesthetic objects are often experienced in sets.

Evaluations of critical items can either diverge from, or converge towards, evaluations of context items that are relatively more or less extreme on a given dimension (Schwarz & Bless, 2007). For example, Jones and Buckingham (2005) asked women to rate their body esteem after viewing a photo of either an attractive or an unattractive woman. Women with high self-esteem identified with the attractive woman, and thus reported higher ratings of body-esteem after
viewing the attractive (vs. unattractive) photo. This is an example of an assimilation effect, as evaluations of body-esteem were more positive following the positive referent. Women with low self-esteem did not identify with the attractive woman, and thus rated their body-esteem lower after viewing the attractive (vs. unattractive) photo. This is an example of a contrast effect in that evaluations of body-esteem were more negative following the positive referent.

The perceived similarity of context and critical stimuli has been found to determine whether context manipulations produce assimilation or contrast. For example, Dolese et al. (2005) performed a study in which participants provided liking ratings for two hedonically positive critical paintings either in isolation (control condition), or after providing liking ratings for five hedonically negative context paintings. The critical paintings received higher liking ratings, as compared to the control condition, when they were labelled as having been painted by the same artist (i.e., a contrast effect). Liking ratings for the critical paintings did not differ from the control condition or the same artist label condition when the context and critical paintings were labelled as having been painted by different artists.

Dolese et al.’s findings are consistent with Parducci’s (1963; 1965) range-frequency model, which treats evaluative judgments as a process of localizing a set of stimuli within the range of a rating instrument. This model focuses on the effects of two aspects of the evaluative context: The range of stimulus quality (i.e., the difference between the lowest- and highest-quality context stimuli), and the frequency of context stimuli at each level of stimulus quality (e.g., whether the context is on average neutral, low, or high quality).

In contrast, studies of social judgment often report a different relationship between the perceived similarity of context and critical stimuli and the direction of the ensuing context effect. To illustrate, Damisch, Mussweiler, and Plessner (2006) had participants rate a critical average-
quality gymnastics routine after rating either a low- or high-quality context routine. When the context and critical gymnasts were described as being on the same national team, ratings of the critical routine assimilated toward the context routine. Conversely, when the two gymnasts were described as being from different national teams, ratings of the critical routine contrasted from ratings of the context routine.

Damisch et al.’s findings support Mussweiler’s (2003) selective-accessibility model, which outlines comparison processes that can generate either contrast or assimilation effects. When evaluative judgments are made along a particular dimension (e.g., beauty) for a series of stimuli, judgment-relevant information from prior judgments should be readily accessible to serve as a comparison standard for each subsequent judgment. How this information is used depends on the perceived relationship between the context and critical stimuli. When a critical stimulus is perceived to be similar to the context, information about past judgments that is consistent with the critical stimulus is retrieved (i.e., a similarity search). This leads the critical stimulus to seem more like the context, thus eliciting an assimilation effect. When a critical stimulus is perceived to be dissimilar from the context, information about past judgments that is inconsistent with the critical stimulus is retrieved (i.e., a dissimilarity search). This leads the critical stimulus to seem less like the context, thus eliciting a contrast effect.

In Experiment 1, we tested whether assimilation or contrast emerges depending on the degree of thematic similarity between the context photos and the critical photos. The critical photos were of modern buildings, and the theme of the context photos was either the same (modern buildings), similar (older buildings), or different (landscapes). Participants viewed a block of either low- or high-beauty context photos and then rated the beauty of a block of critical
average-beauty photos. Experiment 2 then reports three additional attempts to modulate the direction of the context effects obtained in the same-theme condition in Experiment 1.

According to the range-frequency model, a contrast effect should occur in the same-theme condition, for two reasons. First, in a low-beauty context, the average-beauty photos are relatively more beautiful than the context photos, whereas in a high-beauty context, they are relatively less beautiful than the context photos. Thus, the range of stimuli differs across contexts. Second, the frequency of less versus more beautiful stimuli also differs across contexts. The context block ensures that either the lower end (low-beauty context) or higher end (high-beauty context) of the rating scale is well defined, thus pushing participants’ ratings of critical photos in the opposite direction. Neither contrast nor assimilation are expected in the different-theme condition, because the range-frequency model predicts that contrast effects will occur only when context and critical photos are deemed categorically similar.

According to the selective-accessibility model, an assimilation effect should occur in the same-theme condition. Here, thematic similarities between the context and critical images should lead participants to perceive these two types of stimuli as similar, thus engaging a similarity-based search that should yield assimilation. The selective-accessibility model predicts a contrast effect in the different-theme condition. Here, thematic differences between the context and critical images should lead participants to perceive these two types of stimuli as different, thus engaging a difference-based search that should yield contrast.

In sum, the range-frequency model predicts a contrast effect in the same-theme condition and no context effect in the different theme condition, whereas the selective-accessibility model predicts an assimilation effect and a contrast effect, respectively. The inclusion of the similar-theme condition in Experiment 1 allowed us to determine whether an intermediate level of
similarity between context and critical photos produces context effects that pair with the same-
theme condition or with the different-theme condition—a question that could not be addressed in
previous studies, given that only two levels of similarity (same vs. different) were tested.

Our secondary goal was to measure the duration of context effects. To our knowledge,
previous research has not examined this question, and most models of context effects (including
the range-frequency model) do not explicitly address it (e.g., Mussweiler, 2003; Parducci, 1965;
Schwarz & Bless, 2007). However, Mussweiler’s selective-accessibility model states that for any
given judgment in a series, knowledge that was brought to mind to make the preceding judgment
should be most highly accessible. Thus, this account is most consistent with a short-lasting
influence of context. Damisch et al. (2006) examined performance judgments for Olympic
gymnasts. They found a positive correlation between a performance N and performance N-1, and
a weaker correlation between performance N and performance N-2. If the influence of a context
stimulus on evaluations of a critical stimulus decreases with more intervening stimuli, then the
influence of the context block should be relatively short. The photos in our critical block were
identical across context groups, therefore the two context groups’ recent experiences become
more similar across the critical block. This shared context should eventually eliminate the
context effect in the critical block, and the rate with which this happens should be informative
about the accessibility of the context photos and the duration of their influence.

**Experiment 1: Context beauty affects beauty ratings, but context similarity does not**

Participants viewed a set of either low- or high-beauty photos (context block) before
rating a set of average-beauty photos (critical block). The context and critical photos were either
of the same theme, a similar theme, or a different theme (Figure 1).

**Method**
Participants

Participants were drawn from a research participation pool of University of Calgary undergraduate psychology students. Across all participants in this pool, the mean age was 20.93 years, 72.19% were female, and ethnic background was distributed as follows: 46.03% European, 21.33% East Asian, 13.72% Southeast Asian, and 8.47% mixed ancestry (the remaining participants identified as ‘other’). Two hundred and twenty six (155 female) participants participated in groups of up to 15 for course credit. They were randomly assigned to a context beauty condition (low vs. high) in the same ($N$s = 31 vs. 33), similar ($N$s = 28 vs. 29), or different ($N$s = 54 vs. 51) theme condition.

Stimuli

The stimuli were digital photos (cropped to 800 x 600 pixels) selected from online sources. Photos were used because such stimuli are frequently evaluated for their aesthetic qualities, they could readily be compiled and controlled, and the particular photos were novel to participants. The critical photos depicted modern buildings in all conditions of Experiments 1 and 2. The context photos either depicted modern buildings (same theme), older buildings (similar theme), or landscapes (different theme). Additional participants ($N$ = 240) from the same pool rated a large set of one of the three themes in an online norming study. They were asked to consider the beauty of each photo (not its subject matter) for 5 s before rating it on a 7-point Likert scale (1 = not at all beautiful, 4 = somewhat beautiful, 7 = very beautiful). Mean beauty ratings and consistency of ratings were used to select 20 critical average-beauty photos of modern buildings ($M$ = 4.28), and 20 low-beauty (same theme $M$ = 3.00; similar theme $M$ = 2.97; different theme $M$ = 2.97) and 20 high-beauty (same theme $M$ = 5.36; similar theme $M$ = 5.35; different theme $M$ = 5.41) context photos for each set. It proved difficult to find photos that were
CONTRAST EFFECTS ON BEAUTY RATINGS

consistently rated extremely low or high in beauty. Nonetheless, beauty level differed within each set (low < high; \( ps < .001 \)), and there were no differences between themes for a given beauty level (modern buildings = older buildings = landscapes; \( Fs < 1 \)).

**Procedure**

Photos were projected on a large screen in a small classroom using a laptop and data projector. The projected photos were approximately 1 m wide or tall, depending on whether the photo was in landscape or portrait orientation. Participants were told we were studying how people define beauty. They were asked to consider the overall beauty of each photo in the first block, rather than its subject matter. The 20 context block photos were then presented one at a time, for 10 s each, with a 5 s blank screen between photos. Participants were then asked to rate the beauty of each photo in the critical block using the same 7-point scale described above. The 20 critical block photos were presented in the same manner as the context block, and participants wrote their ratings on a numbered sheet of paper during the 5 s blank screen after each photo.

**Results and Discussion**

The mean beauty ratings for the critical photos (Table 1) were analyzed using a 2 (context beauty: low vs. high) by 3 (context similarity: same vs. similar vs. different) by 2 (critical block half: first vs. second) mixed-factor ANOVA. Critical block half was included because the contrast effect eroded across the critical block, as anticipated. Results were significant at the .05 level unless otherwise noted. Eta-squared \((\eta^2)\) is reported as a measure of effect size.

There was a main effect of context beauty: critical block photos were deemed more beautiful after a low-beauty context than a high-beauty context (4.42 vs. 4.17), \( F(1, 220) = 5.88, MSE = 6.87, \eta^2 = .03 \). Thus, ratings of the critical photos contrasted from the context photos
rather than being assimilated toward them. A main effect of critical block half reflected a
decrease in ratings across block halves (4.36 vs. 4.23), \( F(1, 220) = 4.99, \text{MSE} = 1.46, \eta^2 = .02 \).
Context beauty and critical block half interacted, \( F(2, 220) = 30.23, \text{MSE} = 8.87, \eta^2 = 12 \); the
contrast effect was present in the first half of the critical block (4.61 vs. 4.10), \( F(1, 220) = 27.09, \text{MSE} = 15.68, \eta^2 = .11 \), but absent in the second (4.21 vs. 4.25), \( F < 1 \). The main effect of context similarity was not significant, and context similarity did not interact with context beauty or
critical block half, \( Fs < 1 \). The three-way interaction was not significant, \( F(2, 220) = 2.28, \text{MSE} = 0.67, p = .11, \eta^2 = .01 \). The contrast effect in the first critical block half was significant within
each theme condition \( (ps < .01) \).

**Experiment 2: Can we “knock down” context effects on beauty ratings?**

In Experiment 1, the first half of the critical block photos were deemed more beautiful
after a low-beauty (vs. high-beauty) set of context photos. Experiment 2 examined the boundary
conditions for obtaining a contrast effect in the same-theme condition, and also provided an
additional test of the range-frequency and selective-accessibility models.

**Method**

**Participants**

Two hundred and one (145 female) additional participants from the same pool as
Experiment 1 were randomly assigned to a low- vs. high-beauty context in an ignore \( (Ns = 26 \text{ vs. } 29) \), count \( (Ns = 25 \text{ vs. } 24) \), or frame \( (Ns = 48 \text{ vs. } 49) \) condition.

**Procedure**

Experiment 2 was identical to the same-theme condition in Experiment 1 with the
following changes. In the *ignore* condition, after the context block participants were told that the
photos they just viewed had been presented to try to bias their evaluations of the critical photos.
Participants were instructed to try to not let their experiences with the context photos influence their ratings of the critical photos. This condition was conducted to determine whether contrast effects on aesthetic evaluations can be intentionally avoided. In the count condition, participants were asked to count and record the number of windows and floors in the most prominent building in each photo, to test whether the contrast effect persists when participants are not explicitly asked to consider the beauty of the context photos. If so, this would suggest that participants chronically evaluate the beauty of aesthetic stimuli, even when those stimuli are processed for a non-aesthetic purpose. Thus, these two conditions explored the boundary conditions for obtaining contrast effects in our paradigm.

In the frame condition, we attempted to maximize participants’ perceived similarity of the context and critical photo sets and thus to strongly encourage similarity processing. Here, all of the photos in both blocks were presented inside the same distinctive image of a picture frame and participants were told that the same photographer had taken all the photos in both blocks. This condition provided a further test of the range-frequency and selective-accessibility models. The range-frequency model predicts a contrast effect, and the selective-accessibility model predicts an assimilation effect.

**Results and Discussion**

The mean beauty ratings for the critical photos (see Table 1) were analyzed as in Experiment 1. The contrast effect on beauty ratings persisted under the new manipulations; critical photos were deemed more beautiful after a low-beauty context than after a high-beauty context (4.50 vs. 4.14), $F(1, 160) = 13.65, MSE = 10.61, \eta^2 = .08$. Beauty ratings did not decrease across critical block halves (4.32 vs. 4.32), $F < 1$. However, context beauty and critical block half interacted, $F(1, 160) = 22.98, MSE = 5.11, \eta^2 = .13$, reflecting a contrast effect in the
first half (4.63 vs. 4.03), $F(1, 160) = 34.97$, $MSE = 15.21$, $\eta^2 = .18$, but not in the second half of the critical block (4.38 vs. 4.28), $F < 1$. There was no main effect of instruction condition, $F < 1$, nor did instruction condition interact with context beauty, $F < 1$, critical block half, $F(2, 160) = 1.54$, $MSE = 0.34$, $p = .21$, $\eta^2 = .01$, or both factors, $F(2, 160) = 1.29$, $MSE = 0.28$, $p = .27$, $\eta^2 = .01$. As in Experiment 1, the contrast effect in the first critical block half was significant in each condition ($p < .01$). Thus, a contrast effect persisted when participants were asked to avoid the influence of context, when the context photos were processed in a non-aesthetic task, and when we attempted to make the context and critical photos seem more similar than in Experiment 1.

Experiments 1 and 2 establish a contrast effect on beauty ratings that dissipates across the critical block. We pooled the data from Experiments 1 and 2 for a more detailed analysis of the persistence of the contrast effect, admittedly confounding decay (time delay) and interference (number of intervening items). Figure 2 plots the mean beauty ratings for pairs of critical block trials within each context condition. As indicated by the 95% confidence intervals in Figure 2, the contrast effect was significant throughout the first half of the critical block, and was absent throughout most of the second half of the critical block. Thus, we found that the influence of exposure to a set of context stimuli persisted for ten trials, but was gradually eroded by the shared stimuli experienced by both context-beauty groups during the critical block.

**General Discussion**

Context manipulations have been found to affect subjective evaluations of critical items, whether the context and critical items are conceptually similar (e.g., Zellner, Allen, Henley & Parker, 2006) or different (e.g., Kim & Meyers-Levy, 2008). We obtained a contrast effect on beauty ratings of photos regardless of the similarity of the context and critical photos. This pattern is surprising, given that the perceived similarity between context and critical stimuli often
dictates whether assimilation or contrast occurs (e.g., Damisch et al., 2006; Dolese et al., 2005). Dolese et al. argued, based on Fechner’s principle of aesthetic contrast (as cited in Beebe-Center, 1932/1965), that comparative aesthetic judgments should yield contrast effects if the context and critical stimuli are perceived to be similar. Conversely, the categorization of context and critical stimuli as dissimilar can eliminate contrast effects on hedonic evaluations (e.g., Rota & Zellner, 2007) or at least reduce their magnitude (e.g., Dolese et al., 2005; Zellner, Kern, & Parker, 2002; Zellner, Rohm, Bassetti, & Parker, 2003). Although landscapes and modern buildings are nominally different, our participants may have deemed them to be of the same general category (e.g., “photos”). A stronger manipulation might lead participants to perceive the context and critical photos as belonging to different categories, thus reducing or eliminating the contrast effect, consistent with these latter studies.

According to the selective-accessibility model, if our different-theme photos and critical photos were deemed to be from the same category (e.g., “photos”), an assimilation effect should have occurred. In addition, the selective-accessibility model predicts assimilation effects in the same-theme condition. A possible explanation for the absence of assimilation effects in our studies is Mussweiler’s (2003) contention that subjective judgments (which would include beauty ratings) generally lead to contrast effects. Subjective judgments, as opposed to objective judgments such as distance or weight, are not externally anchored—participants must calibrate their judgments within the range of the rating instrument as the rating task progresses. Therefore, the context stimuli, regardless of their thematic type, may have anchored participants’ ratings at one end of the rating scale, thus displacing ratings of the critical stimuli in the opposite direction.

An alternative possibility is that the influence of context similarity on contrast effects may be modulated by the experimental design. Some studies have participants evaluate a series
of context stimuli before evaluating a series of critical stimuli (blocked design), whereas others have participants alternate between evaluating individual context and critical stimuli (mixed design). Blocked designs generally yield contrast effects when the context and critical stimuli are perceived to be similar, and yield no context effects when the two are perceived to be different (e.g., Dolese et al., 2005; Parker et al., 2008). Conversely, mixed designs generally yield assimilation effects when the context and critical stimuli are perceived to be similar, and yield contrast effects when they are perceived to be different (e.g., Damisch et al., 2006; Geiselman et al., 1984). Confirming this dependency, and understanding why it occurs, are important matters for future research.

Our other surprising finding was that we could not “knock down” our same-theme contrast effect in Experiment 2. Presenting the photos in a common picture frame across the two blocks, and telling participants that the photos were taken by the same photographer did not elicit a similarity-seeking strategy that would be expected to produce assimilation (Mussweiler, 2003). Moreover, asking participants to not be influenced by the context photos failed to diminish the effect, despite prior evidence that contextual factors people perceive as influencing their target evaluations tend to elicit a compensation strategy (Schwarz & Bless, 2007). And thirdly, although evaluative outcomes are often influenced by whether artworks are approached for aesthetic or identification purposes (Cupchik, Vartanian, Crawley, & Mikulis, 2009), we found that presenting context photos in a non-aesthetic counting task did not disrupt the contrast effect. It appears that contrast effects on beauty ratings of photos are rather persistent.

Although contrast effects persisted despite our attempts to knock them down, they consistently eroded across the shared critical block as the overlap between the two groups’ recent aesthetic experiences increased. To our knowledge, our finding that the context block (for which,
we add, participants did not make any explicit ratings) modulated beauty ratings across several critical photos is novel. This result will need to be addressed explicitly by accounts of context effects that are relatively silent on the question of persistence (e.g., Mussweiler, 2003; Schwarz & Bless, 2007). Indeed, the vast majority of context effect studies interleave context and critical trials (e.g., Geiselman, Haight, & Kimata, 1984), use only one context and critical trial (e.g., Damisch et al., 2006), or use much shorter blocks of context and critical items (e.g., Cogan, Parker, & Zellner, 2013). Our paradigm thus appears fruitful for exploring the persistence of context effects under various conditions.

Finally, the consistency of the influence of context beauty on aesthetic evaluations also suggests that models of aesthetic experience need to include a role for context beauty. Leder, Belke, Oeberst, and Augustin (2004) identified context as a key factor in aesthetic processing, but they limited its definition to the physical environment in which a stimulus is evaluated. Although this form of context is important (e.g., Cupchik et al., 2009), our findings suggest that theories will need to incorporate a broader definition of context that includes recent experiences with other stimuli.
References


Acknowledgments

The Natural Sciences and Engineering Research Council of Canada supported this research through a Postgraduate Scholarship to CT and a Discovery Grant to GEB. We thank Debra Zellner and an anonymous reviewer for their feedback. Address correspondence to Glen Bodner, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary AB T2N 1N4, Canada, e-mail: bodner@ucalgary.ca.
**Figure 1.** A critical block of average-beauty photos of modern buildings was rated after viewing a block of low- or high-beauty photos of modern buildings (same-theme condition), older buildings (similar-theme condition), or landscapes (different-theme condition).
Figure 2. Mean beauty ratings for critical average-beauty photos following low-beauty and high-beauty photo contexts, for each pair of critical block trials, pooled across Experiments 1 and 2. The 95% between-subjects confidence interval is plotted for each critical block trial pair.
Table 1

*Mean (SE) beauty ratings for each half of the critical average-beauty photo block, after either a low- or high-beauty photo context block, for each condition in Experiments 1 and 2.*

<table>
<thead>
<tr>
<th>Exp/Condition</th>
<th>First half of critical block</th>
<th>Second half of critical block</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-beauty context</td>
<td>High-beauty context</td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same theme</td>
<td>4.61 (0.14)</td>
<td>4.03 (0.13)</td>
</tr>
<tr>
<td>Similar theme</td>
<td>4.79 (0.14)</td>
<td>4.11 (0.15)</td>
</tr>
<tr>
<td>Different theme</td>
<td>4.56 (0.14)</td>
<td>4.12 (0.14)</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ignore</td>
<td>4.61 (0.13)</td>
<td>4.05 (0.12)</td>
</tr>
<tr>
<td>Frame</td>
<td>4.67 (0.12)</td>
<td>4.08 (0.12)</td>
</tr>
<tr>
<td>Count</td>
<td>4.62 (0.13)</td>
<td>3.94 (0.14)</td>
</tr>
</tbody>
</table>