

Affective Color Scales for Colormap Data Visualizations

Halle C. Braun, Kushin Mukherjee , Seth R. Gorelik , and Karen B. Schloss 

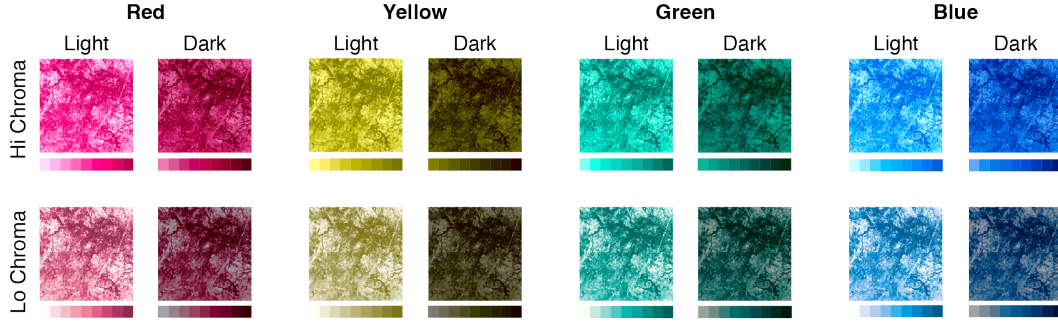


Fig. 1: Example colormaps from Exp. 1, which represent raster data of aboveground live woody biomass [23]. The color scales shown below each colormap were generated in Color Crafter [72] by adjusting hue (red, yellow, green, and blue), lightness (light or dark), and chroma (high chroma or low chroma) within the tool. Here, all 16 color scales were applied to the same underlying dataset, but in the experiment, participants saw each color scale applied to a unique dataset (balanced across participants).

Abstract— Research on affective visualization design has shown that color is an especially powerful feature for influencing the emotional connotation of visualizations. Associations between colors and emotions are largely driven by lightness (e.g., lighter colors are associated with positive emotions, whereas darker colors are associated with negative emotions). Designing visualizations to have all light or all dark colors to convey particular emotions may work well for visualizations in which colors represent categories and spatial channels encode data values. However, this approach poses a problem for visualizations that use color to represent spatial patterns in data (e.g., colormap data visualizations) because lightness contrast is needed to reveal fine details in spatial structure. In this study, we found it is possible to design colormaps that have strong lightness contrast to support spatial vision while communicating clear affective connotation. We also found that affective connotation depended not only on the color scales used to construct the colormaps, but also the frequency with which colors appeared in the map, as determined by the underlying dataset (data-dependence hypothesis). These results emphasize the importance of data-aware design, which accounts for not only the design features that encode data (e.g., colors, shapes, textures), but also how those design features are instantiated in a visualization, given the properties of the data.

Index Terms—Visual reasoning, visual communication, color cognition, affective science, emotion, scalar field, data-aware design

1 INTRODUCTION

In colormap data visualizations, gradations of magnitude are mapped onto gradations of color to illustrate spatial patterns within a dataset. A large body of research on colormap (a.k.a. heatmap) design has focused on which kinds of color scales (a.k.a. color ramps) are most effective for revealing patterns in data [3, 11, 12, 43, 46, 56–60, 72, 88, 90] and how to assign the endpoints of color scales (e.g., light vs. dark) to endpoints of data magnitude (e.g., low vs. high) to help observers interpret the meaning of colors in colormaps [14, 42, 65, 69, 71, 73]. The aforementioned work focused on people's ability to detect and analyze patterns of data, stemming from the general emphasis in visualization research on accuracy, conciseness, and clarity (c.f. [36]). However, another aspect of visualization design has received a recent surge of attention—*affective*, or emotional, aspects of design [2, 6, 9, 34, 35, 37, 62, 70, 87, 89]—which can influence engagement, recall, and aesthetic response (see Lan et al. [36] for a review). In the domain of color,

previous work has focused on understanding affect of color *palettes* (i.e., discrete colors to visualize categorical data) [2, 6, 9], and explored using palettes from paintings to imbue colormaps with affective connotations [62]. Here, we aimed to understand factors that influence affective connotation of colormaps for visualizing continuous data.

The discussion of affect in visualization is often couched in terms of whether a design is embellished to be 'emotional' or kept minimalistic to be 'neutral' (c.f. [36]), but an alternative lens is that all designs can be situated in a continuous space of affective connotation. After all, seemingly neutral 'minimalistic' grayscale visualizations are reportedly linked to negative emotions [9, 89]. From this perspective, research on affective visualization design aims to understand factors that link emotion and design. These links can take two possible forms: *affective connotations*—associations observers have between the design and emotional concepts or *affective state*—emotions observers feel when they look at visualizations. Evidence suggests that color is among the strongest features linking emotion to visualization, compared to other features such as data patterns (e.g., positive/negative slope) or chart type (e.g., scatterplot, bar chart) [9].

Previous work on links between color and emotion has largely focused on individual colors [13, 16, 32, 75, 80, 82, 83] (see Jonauskaitė and Mohr [28] for a review), color pairs [48, 49], faces/scenes [51, 78], or color palettes for visualizations of categorical data [2, 6, 9], but less so on visualizations of continuous data [62, 89]. A general theme across prior studies is that color-emotion links are largely driven by lightness and chroma. For example, *positive* colors are light and high in chroma, *negative* colors are dark and low in chroma, and *calm* or *less angry* colors are light and low in chroma [6, 15, 16, 26, 68, 84]. Although hue plays a role to some extent (and for some emotions more than others),

- Halle C. Braun is with the University of Wisconsin–Madison. E-mail: hallebraun43@gmail.com
- Kushin Mukherjee is with Stanford University. E-mail: kushinm@stanford.edu
- Seth R. Gorelik is with Woodwell Climate Research Center. E-mail: sgorelik@woodwellclimate.org
- Karen B. Schloss is with the University of Wisconsin–Madison. E-mail: kschloss@wisc.edu

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: [xx.xxx/TVCG.201x.xxxxxx](https://doi.org/10.1109/TVCG.201x.xxxxxx)

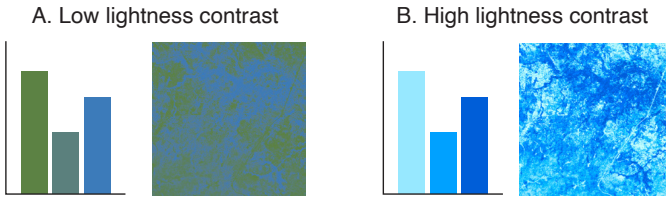


Fig. 2: Visualizations of categorical data (bar chart) and continuous data (colormap) using colors that have (A) low or (B) high lightness contrast.

many hues have similar affective connotation, especially if the colors are light [6, 68].

The strong link between lightness and emotion has interesting implications for visualization design. For visualizations of categorical data (e.g., bar charts), in which data magnitude is encoded using spatial channels (e.g., position, length) and categories are encoded using color, colors of similar lightness may signal strong affective connotation without harming the observer’s ability to find key patterns in the data. For example, comparing the heights of the bars in Fig. 2A (low lightness contrast) seems as easy as in Fig. 2B (high lightness contrast). But, for visualizations of continuous data, in which data magnitude is encoded using dimensions of color, minimizing lightness variability can pose problems because lightness contrast is important for spatial vision. For example, observing spatial structure in the low lightness contrast colormap in Fig. 2A is harder than in the high lightness contrast colormap Fig. 2B. This difficulty is because the visual system relies on variation in lightness to detect spatial structure in data [81], particularly for high spatial frequencies [59] (e.g., fine details in a geographical map). Generally, spatial vision (e.g., contrast sensitivity in high spatial frequencies, orientation sensitivity, pattern and shape recognition) is worse when colors are isoluminant [19, 20, 33, 40, 59]. Yet, introducing lightness variability to support spatial vision will also introduce variability in affective connotation because of the strong link between lightness and emotion. Thus, we asked, **is it possible to produce colormap data visualizations with strong lightness contrast that also have systematic affective connotations?**

We addressed this question by investigating color-emotion associations for colormaps that differ systematically in lightness, chroma, and hue while preserving lightness contrast (Fig. 1). We created the colormaps using color scales generated by Color Crafter [72], a tool that allows for flexible and automatic production of designer-grade color scales (such as ColorBrewer color scales [24]). We then applied the color scales to raster data of aboveground live woody biomass [23] to create colormap data visualizations (Fig. 1). We used a real-world dataset as opposed to synthetic data to support ecological validity of our experiment while avoiding landmarks that would be recognizable in the geographical regions (henceforth, we use ‘subtile’ to refer to these geographical regions). Based on results from a pilot study, we also ensured that the subtiles had similar affective connotation when presented in grayscale so observed effects could be attributed to color and not to spatial aspects of the data. In Exp. 1, the values in the underlying dataset were well-distributed so that each color of a color scale was well-represented in the corresponding colormaps. In Exp. 2, we shifted the underlying data distributions to test if affective connotation depended not only on the colors in the color scale, but also on the relative frequency with which each color appeared in the colormap (data-dependence hypothesis).

Contributions. In this paper we make three main contributions: (1) We provide evidence that colormaps with large lightness variation can have systematic affective connotations. (2) We provide support for the data-dependence hypothesis, which emphasizes the need for data-aware design—affective connotation depends not only on the color scale used to construct a colormap, but also on the relative proportion of colors determined by the underlying dataset. (3) We describe affective connotations of colormaps generated from 16 different color scales and how connotation is modulated by the underlying dataset, which can inform the use of color in future research and visualization design. In the General Discussion, we consider open questions to build on this foundation towards a comprehensive understanding of affective

visualization design for colormaps.

2 BACKGROUND

2.1 Affective visualization design

As defined by Lan et al. [36], *affective visualization design* concerns data visualization designed to communicate and influence emotion. It is part of a larger class of *affective design*, which also includes designs to visualize aspects of emotion (e.g., measures of heart rate) and studies of how priming an observer’s affective state influences their evaluations and interpretations of visualizations. Lan et al. [36] reviewed several kinds of methods in affective visualization design, including (1) ‘sensation’, pertaining to links between perceptual experiences and emotion, (2) ‘narrative’, pertaining to storytelling techniques used to communicate data, (3) ‘behavior’, pertaining to manipulating/operating on data, creating/building data visualizations, walking/wandering through visualizations, and (4) ‘context’, pertaining to the surrounding environment when people view or interact with data. ‘Sensation’ is the most pertinent factor for the present study, so the rest of this section will focus on links between perceptual features and emotions.

When considering links between perceptual features and emotions, there is a qualitative distinction between *affective connotation* (i.e., conceptual association) between perceptual features and emotions, and *affective state* (i.e., feelings) elicited from perceptual features. For example, people strongly associate high chroma reds with the concept of anger, but that does not necessarily mean they *feel* anger when they see high chroma reds [28, 68]. This distinction is perhaps best illustrated in the domain of music. Individuals with *musical anhedonia* do not experience pleasure while listening to music [8, 31, 41], but their ability to identify emotional associations with music is similar to typical listeners [64] (see [8, 31, 41] for potential mechanisms).

Different kinds of methods are used for assessing affective connotation and affective state. Methods for assessing affective connotation are relatively straightforward and include having observers rate associations between perceptual features and emotion-related concepts, match visualizations to emotion-related concepts, or assess which features observers choose when designing visualizations to communicate particular emotion concepts. In contrast, methods for assessing affective state in response to perceptual stimuli are more complicated and researchers have concerns about interpreting the measures [5, 29]. One method is having participants self-report on their emotions through ratings or verbal descriptions, but there are concerns that observers may not distinguish between felt emotions and associations with emotions when self-reporting on affective state and that ratings may be influenced by the wording or images used in emotion scales. A second class of methods involve physiological measures, but physiological measures can contradict self-report measures and it is unclear how to interpret such conflicts. Given complications surrounding the assessment of felt emotions in response to visual features, this study focuses on affective connotation of colormap visualizations.

2.2 Correspondences between color and emotion

Researchers have been studying links between individual colors and emotion for over 100 years (see Jonauskaitė and Mohr [28] for a review). Evidence suggests that associations between single colors and emotion are largely driven by lightness and chroma, with some systematic effects of hue [15, 16, 26, 68, 84]. For example, Schloss et al. [68] asked participants to rate emotional associations for single colors that varied independently in hue angle, lightness, and chroma (similar to our approach in the present study). They found that lighter, higher chroma colors were happier (and less sad) than darker, lower chroma colors. They assessed effects of hue by testing for differences across the red/green and yellow/blue dimension (a^* and b^* in CIELAB space, respectively) and found that bluer colors were happier than yellower colors, particularly for darker colors. This finding may seem surprising given the conventional notion that yellows are happier than blues, but that notion is likely an artifact of the most typical yellows being lighter than the most typical blues (and colors of yellow hue crossing the category boundary from yellow to brown as they darken and are less likely

to be called yellow) [16, 28, 68]. The dominance of lightness in color-emotion associations may explain why red-green dichromatic observers have similar color-emotion associations as observers with typical color vision [27]—both groups can perceive lightness differences.

In their study of affective connotation of color palettes, Bartram et al. [6] also found strong effects of lightness and chroma. They began by generating candidate colors for a design study by analyzing the color palettes of images tagged with eight emotion concepts that span the pleasure (valence) and arousal (activation) axes of the pleasure, arousal, dominance (PAD) model of affect [54, 61]: *calm*, *exciting*, *serious*, *playful*, *positive*, *negative*, *disturbing*, and *trustworthy*. Then, they had users compose palettes for visualizations that best communicated each emotion. Colors selected for the terms *positive*, *playful*, and *calm* were lighter than the colors selected for *disturbing*, *serious*, and *negative*. In terms of chroma, colors selected for *exciting*, *playful*, and *positive* were high in chroma, and colors for *calm* were low in chroma. There were also some effects of hue, such as *calm* being linked to cooler hues (blues and greens) and *exciting* being linked to warmer hues (red, oranges, yellows). Bartram et al. [6] found similar results when they asked observers to rank visualizations for each affective term.

Extending from Bartram et al. [6], Anderson and Robinson [2] investigated the effects of affective congruence between color palette and data topic on interpretations and evaluations of choropleth maps. They chose color palettes and topics that were *positive*, *negative*, *calm*, or *exciting* (palettes from [6]) and paired them so they were congruent (e.g., high chroma colors with positive topics like ‘kinds of desserts’; low chroma colors with negative topics like ‘methods of homicide’) or incongruent (i.e., opposite color/topic pairing). Congruency was defined with respect to the palette as a whole, without considering semantic discriminability of the colors [44, 66] or optimal color-concept assignments [39, 44, 66, 67]. As such, they did not find effects of affective congruence on interpretive tasks that required mapping concepts to individual colors. Yet, they did find that affectively congruent visualizations had amplified emotional connotation (e.g., maps of pleasurable topics were judged as especially pleasurable when the colors were positive rather than negative) and were rated as being more appropriate. These results suggest there are important downstream implications of the affective connotation of color in visualizations.

Most recently, Blair et al. [9] assessed self-reported affective state (felt emotion) in response to color, as well as other aspects of visualization design: chart type, data trend, data variance, and data density. They found that all of these factors contributed to affective state, but color was an especially powerful cue. Moreover, reports of affective state were similar for abstract color palettes (i.e., colored squares) and colors presented in scatterplot visualizations. These findings suggest that affective response to color may be agnostic to presentation format and design may not need to be data-aware, at least for categorical visualizations where all palette colors can have similar affective connotation. We will return to this issue in Exp. 2 and the General Discussion.

Research on links between color and emotion for visualization primarily focused on categorical data, but there are some exceptions. Zhang et al. [89] studied links between emotional response to colormaps and visualization comprehension and found that colormaps eliciting more positive affective responses supported greater comprehension. They did not report on what aspects of color appearance influenced affective response, but their results imply that colormaps containing lightness contrast may evoke specific emotions.

To design affectively connotative colormaps, Samsel et al. [62] took inspiration from colors in paintings. Their process involved choosing a painting that had a desired affective connotation, selecting representative colors from the painting to form a palette, using the palette as input into ColorMoves to create a color scale, and then applying the color scale to the dataset to produce the colormap visualization. Although they did not assess affective connotation in the resulting colormaps, the premise was that the colormaps would share the connotation of paintings from which the colors were sourced.

Building on this prior work, we aimed to understand the factors influencing affective connotation of colormap data visualizations, focusing on the dimensions of color and properties of the underlying dataset.

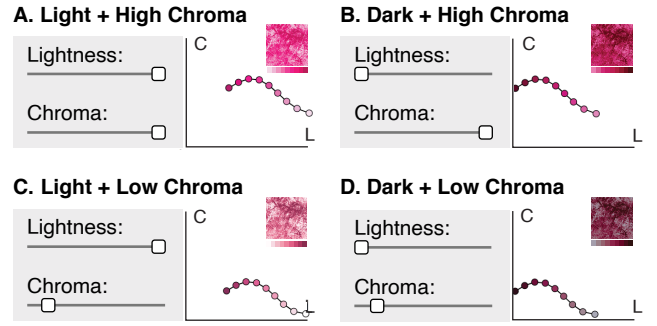


Fig. 3: Illustration of how color scales were generated in Color Crafter by adjusting lightness and chroma sliders (note, in Color Crafter the scale controlling chroma is labeled ‘saturation’). The charts rightward of slider pairs show Color Crafter’s charts of the 9-point color scale plotted in terms of lightness (L) and chroma (C) of CIELCh space. The inset in each plot shows an example colormap produced from that color scale.

3 EXPERIMENT 1

Exp. 1 assessed color-emotion associations for colormaps that differed systematically in hue, lightness, and chroma. The underlying datasets were selected such that all colors in the color scales were well-represented in the colormap visualizations.

We hypothesized that although all the colormaps varied in lightness, their overall lightness level would drive affective connotation. Moreover, based on the findings for single colors [16, 28, 68, 84], we hypothesized that positive affective terms would be associated with lighter, higher chroma, and bluer colormaps, whereas negative affective terms would be associated with darker, lower chroma, and yellower colormaps. We also hypothesized that the concept *calm* would be associated with cooler, lighter, and lower chroma colormaps [6]. The experimental stimuli, data, and analysis scripts for both experiments in this study can be found in our GitHub repository¹.

3.1 Methods

3.1.1 Participants

The participants ($n = 35$) were undergraduates from the University of Wisconsin–Madison (UW–Madison) who received extra credit in their Psychology course (mean age: 18.8, range: 18–23). They provided demographic information in free-response text boxes. Their genders included 16 women and 19 men (no other genders reported). Their race/ethnicity included 27 White, 5 Asian, 2 Hispanic, 1 Black, 1 Arab, 1 African-American, and 1 Middle Eastern (the number of entries exceeds the number of participants because some reported > 1 race/ethnicity). For all experiments in this study, all participants gave informed consent and the UW–Madison IRB approved the protocol.

Our target sample was $n = 32$, established as follows. Based on the sample size of $n = 20$ in Exp. 2 of [68] (analogous to the present experiment but for single colors), we sought to collect data from at least 20 participants. To reach $n \geq 20$ and to accommodate the 16 conditions in the Latin square design that balanced the color scale and dataset assignments across participants (Section 3.1.2), we doubled the Latin square resulting in target $n = 32$. We ran $n = 35$ to accommodate excluding 3 participants with atypical color vision determined by self-report questions or errors on more than 2 of 11 digital Ishihara plates presented at the end of the study (criteria determined in advance).

3.1.2 Design and displays

Based on the approach in Schloss et al. [68] for single colors, we constructed colormaps generated from 16 color scales, comprising the orthogonal combination of 4 hues (red, yellow, green, and blue) $\times$ 2 lightness levels (light, dark) $\times$ 2 chroma levels (low, high) (Fig. 1). We generated the color scales in Color Crafter [72], detailed below. Each participant rated the association between each of 16 colormaps and each of 8 emotion concepts from Bartram et al. [6] (*positive*, *playful*,

¹https://github.com/SchlossVRL/color_scales_affect

exciting, negative, serious, disturbing, calm, and trustworthy) twice, resulting in 256 trials.

In the colormaps in Fig. 1, each color scale was applied to the same underlying dataset but in the experiment, each participant saw 16 colormaps created from a unique pairing of color scales to underlying datasets. That is, they never saw colormaps constructed from the same underlying dataset in more than one color scale. To achieve this experimental design, we produced 256 colormaps (16 color scales $\times$ 16 underlying datasets) and used a Latin square design to balance the assignment of color scales to underlying datasets across participants. This design produced 16 groups, and participants were randomly assigned to group. Within participants, the trials were presented in a randomized block design—participants judged all 16 colormaps in a random order twice within one emotion concept before going on to the next emotion concept (32 trials per block). The order of emotion concepts was also randomized for each participant.

The following sections describe how we generated the color scales and selected the 16 underlying datasets to construct the colormaps.

Generating color scales. We created the 16 color scales using *Color Crafter* [72], which enables inputting a seed color to generate color scales and provides sliders to shift the overall lightness (L^* in CIELCh) and chroma (C^* in CIELCh).

We defined the seed color for each hue in CIE LCh space such that $L^* = 90$, $C^* = 5$, and hue angle (h) was 0, 90, 180, 270 degrees for red, yellow, green, and blue seeds, respectively. We chose these hue angles as approximates of the Hering opponent dimensions (red/green, yellow/blue). We chose these lightness and chroma coordinates after exploring different values in *Color Crafter* and determining which seeds produced color scales that were roughly monochromatic (single hue). We then converted these coordinates to CIE XYZ values with a white point of CIE Illuminant D65, and converted to RGB values using the `xyz2RGB` function in MATLAB (assuming a standard sRGB monitor), and then to hexadecimal codes as input seeds for *Color Crafter* (red: #ecdfe3, yellow: #e7e2d9, green: #d8e5e2, blue: #dee3ec). Given a seed color, *Color Crafter* generates nine color scales. We selected the color scale of the target hue that appeared most monochromatic (i.e., had the least hue variability).

Next, we adjusted the chroma and lightness sliders to create four color scales for each hue (Fig. 3). For *light high chroma* color scales, we increased lightness and chroma sliders to their maxima (Fig. 3A). For *dark high chroma* color scales, we kept the chroma slider at its maximum and we reduced the lightness slider to its minimum (Fig. 3B). To create the *dark low chroma* color scales, we kept the lightness slider at its minimum and reduced chroma as far left as possible while preserving monotonicity in the light region of the curve (Fig. 3D).² For *light low chroma* color scales, we kept the chroma slider where it was and minimized the lightness slider (Fig. 3C).

The color coordinates of the color scales are not perfectly controlled (e.g., L^* and C^* are not matched across the light, high chroma color scales for each hue, see Supplementary Material Figures S.1 and S.2), but we used this approach because *Color Crafter* produces effective color scales for data visualization in practice [72]. We could have designed tightly controlled color scales, but they would likely have been less effective for visualization, thus weakening the translatability of our results to design practice. Our analyses control for this variation by coding predictors according to the average CIELAB/CIELCh coordinates of the colors scales. See details in Section 3.2.1.

Selecting the underlying datasets. We created the colormaps by applying the 16 color scales described above to a global raster dataset of aboveground live woody biomass (AGB) [23]. Each pixel in the dataset has a spatial resolution of 30×30 meters and represents megagrams of AGB per hectare (Mg ha^{-1}) in the year 2000. The dataset includes 280 ‘tiles’ across the world ($40,000 \times 40,000$ pixels per tile). We split

²Sliding chroma all the way left (lowering chroma) can cause the chroma of the curve to ‘dip’ in the light range because the color scale crossed zero in the A/B plane of CIELAB space, thereby adding a new hue into the color scale. Aiming to produce roughly monochromatic color scales, we avoided this issue by moving the chroma slider as far left as possible before the ‘dip’.

each tile into smaller $1,000 \times 1,000$ pixel ‘subtiles’ (448,000 subtiles total). We then excluded subtiles that had N/A values, corresponding to 0 Mg ha^{-1} , to avoid including subtiles that had easily recognizable geographical features (e.g., rivers or croplands).

Of the remaining set, we sought to find subtiles that had the most uniform distribution of values, such that when we applied a color scale to produce a colormap, all colors in the color scale would be well-represented in the map. Although we could have generated synthetic datasets that were perfectly uniform, we used real-world data for ecological validity. We used the following steps to filter the subtiles. First, we rescaled the values from units of biomass density to a standardized scale from 0 to 100. Second, we eliminated subtiles whose medians were outside of a range of 40–60 to avoid skewed distributions. Third, for each of the subtiles we grouped the values into nine bins corresponding to the nine colors within the *Color Crafter* color scales and computed the pixel count for each bin. We kept subtiles for which the pixel counts of each of the nine bins fell within an acceptable range of deviation from the expected pixel count for each bin under a uniform distribution (1000×1000 pixels/9 bins).³ This method resulted in 44 maps (Fig. S.3; see CSV file in our [GitHub](#) repository for pixel values).

Identifying subtiles with similar affective connotation. Given our aim to test the influence of color on affective connotation of colormaps, we sought to use maps that all had similar affective connotation when presented in grayscale. Thus, any differences observed in our main experiments would likely be due to color and not spatial structure of the maps (e.g., spatial distribution of pixels). We note that such effects were also mitigated by balancing the assignment of color scale to underlying datasets across participants.

To find grayscale maps with similar affective connotation, we conducted a pilot study in which 46 participants rated the association between each of the 44 maps (Fig. S.3) and each of the eight affective terms tested in Exp. 1 (see Section S.1 for participant demographics). The experimental design and procedure were the same as in Exp. 1, replacing the 16 chromatic colormaps with 44 grayscale colormaps, each rated once.

To determine the most similar maps in terms of affective connotations, we first correlated the mean associations for each pair of maps across the 8 emotion concepts (averaged over participants), which produced a symmetric 44×44 correlation matrix (Fig. S.4A). Hierarchical clustering revealed a central cluster of 26 strongly correlated maps (i.e., similar affective connotation). We selected the 16 maps that had the highest mean correlation with each of the other maps. Examining the affective connotation of these 16 maps in grayscale (Fig. S.4B), they happened to be most associated with *disturbing* and *negative*.

Applying color scales to subtiles. With the final set of 16 subtiles, we applied the 16 color scales to create the colormaps for this experiment. This process produced 256 colormaps for Exp. 1.

Experiment displays. The experiment was administered in participants’ browsers using *jsPsych* [17, 18]. As such, the precise colors each participant saw depended on properties of their own monitors and all reported CIE values are approximations using standard assumptions about RGB displays. This approach is typical in visualization research, which aim to produce results that are robust to variations across displays [21, 44, 74, 76].

During the experimental task, the participants saw displays each containing one of the colormaps centered on the screen scaled to be $200 \text{ px} \times 200 \text{ px}$. This size corresponded to $3.8 \text{ cm} \times 3.8 \text{ cm}$ on a display that was $29.5 \text{ cm} \times 16.5 \text{ cm}$ with a resolution of $1920 \text{ px} \times 1080 \text{ px}$. Below was a slider scale with the endpoints labeled ‘not at all’ and ‘very much’ and an unlabeled midpoint. The background of the monitor was gray (rgb: [128,128,128]).

3.1.3 Procedure

The participants were instructed that they would be presented with a series of colormaps, one at a time, and their task was to rate how much

³We operationalized ‘acceptable’ as within $\pm 1/3$ of the expected bin counts. We explored other metrics to assess deviation from a uniform distribution, but we found this approach was most effective for ensuring that no bin had an outsized number of pixels.

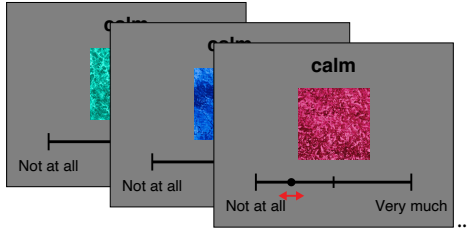


Fig. 4: Example trials for rating color-emotion associations for colormaps.

they associated each map with each of the eight affective concepts on a scale from ‘not at all’ to ‘very much’ (Fig. 4). Before beginning the task, they were shown a list of all eight concepts and all 16 colormaps they would be asked to judge during the experiment. The colormaps were shown in two rows of eight maps and were randomly ordered for each participant. Below the maps was an example rating scale. The participants were asked to look at the concepts and maps and consider which map they associated ‘very much’ and ‘not at all’ with each concept. This ‘anchoring’ task was done so participants would know what associating ‘not at all’ and ‘very much’ meant for them for the colors and concepts in the context of this experiment [50]. The participants were also asked to use the full range of the scale when making their ratings and were instructed that if they somewhat associated a map with a concept, then they should click partway between the middle point and right side of the scale. Participants were informed that they would be asked to rate each map for a given concept before moving on to the next concept. They were also told to make ratings based on their first intuition because we were interested in their initial impressions of each map for a given concept.

The participants were then instructed on how to use the response scale by sliding the cursor to the position where they wanted to make their rating and then letting it go to record their response. They were given four training trials, asking them to move the slider to each endpoint and halfway between the center and each endpoint. If they did not place the slider near the instructed location, they were asked to try again until the correct location was selected. Before each block, participants were told the concept they would be asked to judge for that block. They were then presented with the 32 trials for that block, each separated by 250ms inter-trial interval.

3.2 Results and Discussion

We first analyzed the results in terms of discrete emotions (Fig. 5) and then examined the underlying dimensional structure (Fig. 7A). All analyses were conducted using R version 4.2.1 [55].

3.2.1 Analysis of discrete emotions

Fig. 5 shows the mean association ratings for each of the eight emotion concepts and each of the 16 color scales. We analyzed the pattern of data for each concept using linear mixed-effect regression (LMER) models [4, 7], analogous to the models in [68], to predict participants’ association ratings between the colormaps and that concept. We included fixed effects for each color scale’s mean lightness (L), chroma (C), a^* (red/green), and b^* (yellow/blue) from CIELAB and CIELCh color spaces. We also included interactions for each opponent hue pair ($L \times C \times a^*$ and $L \times C \times b^*$) and pairwise interactions therein following [68], and included by-subject random slopes and intercepts for each fixed effect. We note that averaging a^* or b^* would be nonsensical if the color scales were not roughly monochromatic (Fig. S.1) (e.g., averaging opposite hues along a^* or b^* would cancel to gray). We z-scored each predictor to put them on the same scale so that the resulting model weights could be comparable across predictors.

Given that we fit eight models, one for each emotion, we applied a Bonferroni correction at the model level to correct for multiple comparisons (critical α of $.05/8 = .00625$). Fig. 6A highlights key findings and the full model results are shown in Supplementary Material Table S.4. The dots in Fig. 6A show the sign (+/-) and magnitude (diameter) of the beta weights for each effect that was significant after correction

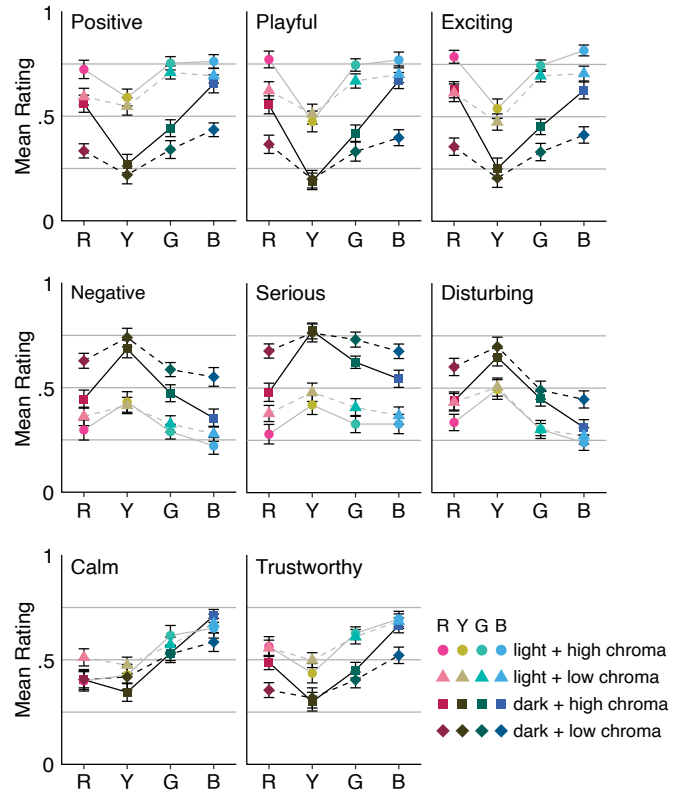


Fig. 5: Mean association ratings for colormaps created from each of the 16 color scales and the eight affective concepts. The x-axis represents hue (Red, Yellow, Green, Blue) and the separate lines represent lightness (dark: dark gray, light: light gray) and chroma (high chroma: solid, low chroma: dashed). The mark color represents the middle color of the color scale used to construct the corresponding colormaps. Error bars represent +/- standard errors of the means (SEM).

(*ns* indicates not significant). None of the 3-way interactions were significant so they were excluded from Fig. 6A.

Generally, the pattern of associations for *playful* and *exciting* were similar to *positive* (Fig. 5, row 1), being strongly associated with lighter, higher chroma, and bluer colors. These emotions also all shared a chroma $\times$ red/green interaction—higher chroma colormaps were especially more *positive*, *playful*, and *exciting* for redder colormaps than greener colormaps. However, the three emotions differed in terms of their other interactions. For *positive* and *exciting*, a chroma $\times$ yellow/blue interaction indicated that higher chroma colors were especially more *positive* and *exciting* when the colors were bluer. For *exciting*, lightness interacted with yellow/blue, indicating that darker colors were especially less exciting when they were yellower.

The pattern of associations was also similar among *negative*, *serious*, and *disturbing* (Fig. 5, row 2), being associated with darker, lower chroma colors. However, whereas *negative* was associated with yellower colors with no effect of red/green, *serious* had no effect of yellow/blue, and *disturbing* was associated with redder colors. These three emotions also differed in terms of which interactions were significant. For *negative* only, a chroma $\times$ lightness interaction indicated that the degree that low chroma colors were more negative was greater when the colormaps were darker. For *negative* and *disturbing*, a chroma $\times$ yellow/blue interaction indicated that increasing chroma especially reduced negativity/disturbingness when colors were bluer. For *serious* and *disturbing*, chroma interacted with red/green, though the patterns were different. For *serious*, greener colors were especially more serious when chroma was higher, whereas for *disturbing*, redder colors were especially more disturbing when chroma was lower.

Finally, *trustworthy* was similar to *calm* (Fig. 5, row 3) in that both were associated with bluer colors with no effect of chroma, but they differed in that *trustworthy* was associated with higher lightness.

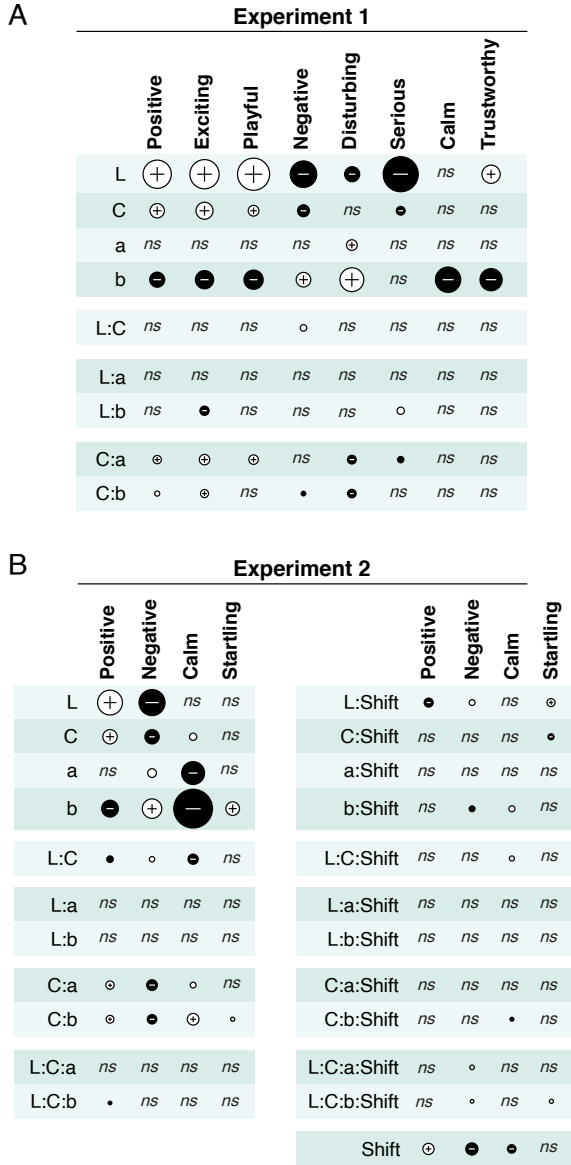


Fig. 6: Summary of LMER model results for (A) Exp. 1 and (B) Exp. 2 (see Supplementary Material Tables S.4 and S.6 for the full output). The factors included lightness (L*), chroma (C*), red/green (a*), yellow/blue (b*), and interactions therein. Dots with +/- symbols indicate significant effects with positive/negative beta weights, respectively ('ns' indicates not significant). Dot diameter is proportional to beta weight to represent effect size (diameter minimum was set to 1.5 pixels to ensure dot visibility).

Together, these results answer our core research question: it is indeed possible to produce colormap data visualizations with strong lightness variation that also have systematic affective connotation.

3.2.2 Analysis of underlying dimensional structure

To understand the nature of the underlying color-affect space, we conducted principal component analysis (PCA) on the mean ratings for colormaps constructed from each color scale for each affective term. This space is well-characterized by a 2D solution, illustrated in Fig. 7A. Each of the 16 points in this space represents each of the 16 color scales (combined over colormaps produced from that color scale). The color scale loadings on PC1 were positively correlated with how *positive*, *playful*, and *exciting* the colors were ($r(14) = 0.98, 0.98, 0.97$, all $ps < 0.001$, respectively) and negatively correlated with how *negative*, *serious*, and *disturbing* the colors were ($r(14) = -0.99, -0.91, -0.97$, all $ps < 0.001$, respectively). As such, PC1 seems to represent *valence* [54] so we labeled it positive vs. negative. PC2 was most strongly

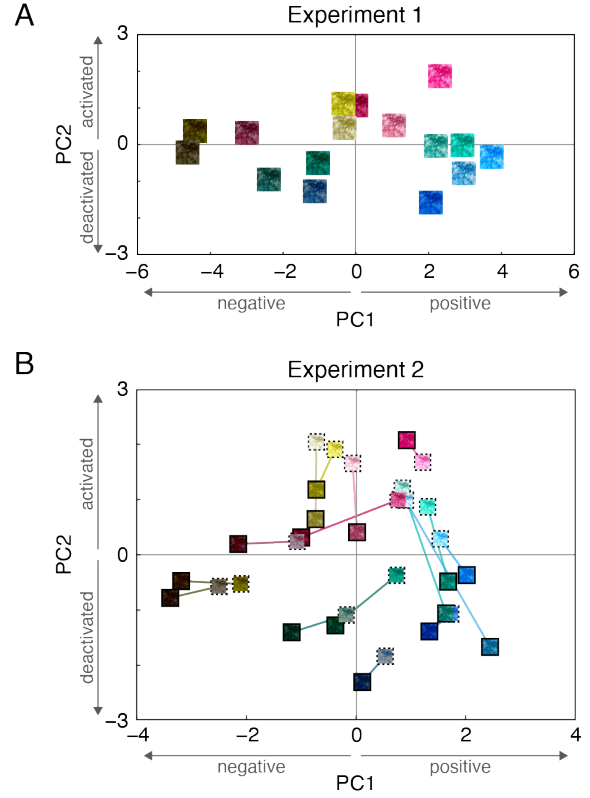


Fig. 7: Results of the PCA conducted on association data from (A) Exp. 1 and (B) Exp. 2. In (A), each of the 16 colormaps represents all colormaps produced using the color scale shown in that map. In (B), the 16 colormaps with solid/dashed outlines represent dark/light shifted colormaps, respectively. Lines connect dark and light shifted colormaps generated from the same color scale.

correlated with how *calm* the colors were ($r(14) = -0.73$, $p < 0.001$) and thus seems to represent arousal/activation [54] so we labeled it activated vs. deactivated. The full set of correlations are in Supplementary Material Fig. S.6. The dimensions of this affective space are similar to dimensions for single colors [1, 28, 50, 80] and color palettes [6].⁴

To characterize the color variation along each PC, we used multiple linear regression with color scale coordinates along the PC as a dependent measure and mean L*, C*, a*, and b* values of each scale as predictors. PC1 (valence) varied from darker, lower chroma, and yellower (negative) to lighter, higher chroma, and bluer (positive) (adjusted model $R^2 = 0.924$). PC2 (activation) varied from greener, bluer, and darker (less activated) to redder, yellower, and lighter (more activated) (adjusted model $R^2 = 0.836$). The full set of results are in Supplementary Material Table S.5.

Summary. Exp. 1 showed that colormaps can have clear affective connotation while preserving the lightness contrast that is important for fine-detail spatial vision. These associations were primarily driven by overall lightness, as well as chroma and yellowness/blueness of the colormaps. Analysis of emotion dimensions showed that higher valence corresponded to lighter, higher chroma, and bluer colormaps, and higher activation corresponded to lighter, yellower, and redder colormaps. Designers can leverage our findings for discrete emotions and reference our affective space when selecting color scales to convey desired emotions for colormap data visualizations.

4 EXPERIMENT 2

In Exp. 1, the values of the subtile were well-distributed over the range, which produced colormap for which all colors in the color scale

⁴The dimensions pleasure, arousal, and dominance (PAD) for emotion are analogous to evaluation, potency, and activity (EPA) for semantic associations [80]. These citations include studies reporting PA or EA as key dimensions.

were well-represented. However, in real-world maps, values can be biased toward one endpoint of the range (e.g., larger or smaller values), resulting in colors that are biased toward one end of the color scale (e.g., darker or lighter colors). Exp. 2 investigated whether affective connotation of colormaps depends simply on the color scales used to construct the colormaps (null; data-agnostic hypothesis), or rather depends on the proportion in which each color appears in the colormap, determined by the underlying dataset (data-dependence hypothesis).

4.1 Methods

4.1.1 Participants

The participants ($n = 34$) were UW–Madison undergraduates who received extra credit in their Psychology course (mean age: 18.9, range: 18–21). They provided demographic information in free-response text boxes. Their genders included 27 women and 7 men (no other genders reported). Their race/ethnicity included 24 White, 5 Asian, 4 Hispanic, 2 Latino/Latina, 1 Middle Eastern, 1 Native American, and 1 Hmong (the number of entries exceeded the number of participants because some participants indicated > 1 race/ethnicity). We excluded 1 participant who indicated atypical color vision (assessed as in Exp. 1).

4.1.2 Design, displays, and procedure

The methods were the same as Exp. 1 with the following exceptions:

Shifted colormaps. Participants judged two versions of colormaps generated from each color scale, a light and dark shifted version (16 color scales $\times$ 2 shift conditions = 32 maps). Fig. 8A illustrates light and dark shifting, starting from an original colormap from Exp. 1. The histogram next to each colormap shows the number of pixels mapped to each color of the color scale (bar color corresponds to color in the color scale). In the light shifted maps, the majority of pixels are in the range of values corresponding to light colors in the color scale, whereas the opposite is true in the dark shifted maps.

To create the shifted versions, we first normalized the values in each subtitle (0–1; min-max normalization). Next, we created ‘light shifted’ maps (i.e., a greater proportion of lighter values) by applying the following formula to each pixel x in the data matrix

$$b - \frac{(b-x)^p}{(b-a)^{p-1}}$$

where a and b were the limits of allowable data values (0 and 1, respectively). The parameter p determined the degree of shift (larger values of p resulting in larger shifts).

The analogous equation for making ‘dark shifted’ colormaps was

$$\frac{(x-a)^p}{(b-a)^{p-1}} + a,$$

We set the value of p to be 4 after visually inspecting the light and dark shifted maps for a range of parameter values. This parameter value resulted in shifted maps that were perceptually distinct from the non-shifted version while preserving enough visual detail that they still appeared to be generated from the same data distribution.

Fig. 8B shows light and dark shifted maps for each of the 16 color scales. For a given participant, the light and dark shifted maps for a given color scale started from the same underlying dataset, and each color scale had a unique dataset (as in Exp. 1).

Emotion concepts. We reduced the set of emotion concepts because associations were highly correlated among concepts in Exp. 1 (Fig. S.5). We retained *positive* and *negative* to span valence and *calm* as an endpoint of activation. Typically, exciting is considered opposite of calm [6,54] but that was not the case in Exp. 1 (Supplementary Material Figs. S.5 and S.7). Thus, we sought a different opposite of *calm* using a natural language processing (NLP) word embedding model (GloVe [52]) and arrived at *startling* (see Supplementary Material Section S.2 for details).

During the experiment, the participants rated their association between each of the 32 colormaps twice for each of the four emotions in a randomized block design as in Exp. 1 (256 trials in total). In the task instructions, we showed participants the full set of concepts for

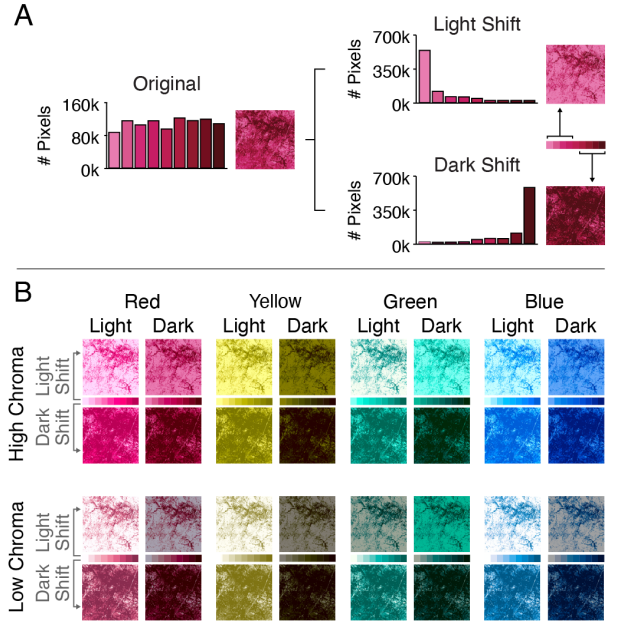


Fig. 8: (A) Diagram showing how the underlying dataset for a colormap from Exp. 1 was light shifted and dark shifted to produce the colormaps in Exp. 2. The histograms for each colormap show the number of pixels in each color from the color scale used to produce the colormap (the bar colors correspond to pixel colors in the colormap). (B) Example colormaps that were created from the same 16 color scales from Exp. 1 (Fig. 1) applied to light shifted or dark shifted datasets.

anchoring (including those from Exp. 1) so they would have the same affective context for performing the task. But, during the task, they only judged *positive*, *negative*, *calm*, and *startling*.

4.2 Results and Discussion

As in Exp. 1, we analyzed the results in terms of discrete emotions (Fig. 9) and then examined the underlying dimensional structure (Fig. 7B).

4.2.1 Analysis of discrete emotions

Fig. 9 shows the mean association ratings for each of the four affective concepts and each of the 16 color scales, depending on whether the underlying datasets were light or dark shifted.

We first analyzed the data using the same LMER models as Exp. 1 but included shift as a new fixed factor (+1 = light shift, -1 = dark shift) and tested for interactions between shift and the other factors. We corrected for multiple comparisons at the model level as in Exp. 1, but here there are four models instead of eight (critical α of $0.05/4 = 0.0125$). Also here, effects combined over shift included 4 judgments (2 shift levels $\times$ 2 repetitions), as opposed to 2 judgments in Exp. 1. With a less strict α and more data per color scale to produce more stable measures, some factors that did not reach significance in Exp. 1 may be considered significant here. The full output of the models are shown in Supplementary Material Table S.6 and key results are summarized in Fig. 6B. The effects of color dimensions combined over shift are shown in Fig. 6B (left), and can be compared to Exp. 1 (Fig. 6A). The effects involving shift are in Fig. 6B (right). We first compare the results combined over shift to the results of Exp. 1, and then focus on effects involving shift to test the data-dependence hypothesis.

Effects combined over shift conditions. As in Exp. 1, *positive* was associated with lighter, higher chroma, and bluer colormaps, and chroma interacted with red/green and yellow/blue. Unlike Exp. 1, lightness $\times$ chroma and lightness $\times$ chroma $\times$ yellow/blue interactions indicated that the increased association for higher chroma colors with *positive* was more extreme when colors were darker, especially when they were also bluer (Fig. 9). Although these effects were not significant in Exp. 1, the pattern was similar (Fig. 5) and the weights were in the same direction (Table S.4). These effects are also analogous to previous results for single colors when rated in terms of *happiness* [68].

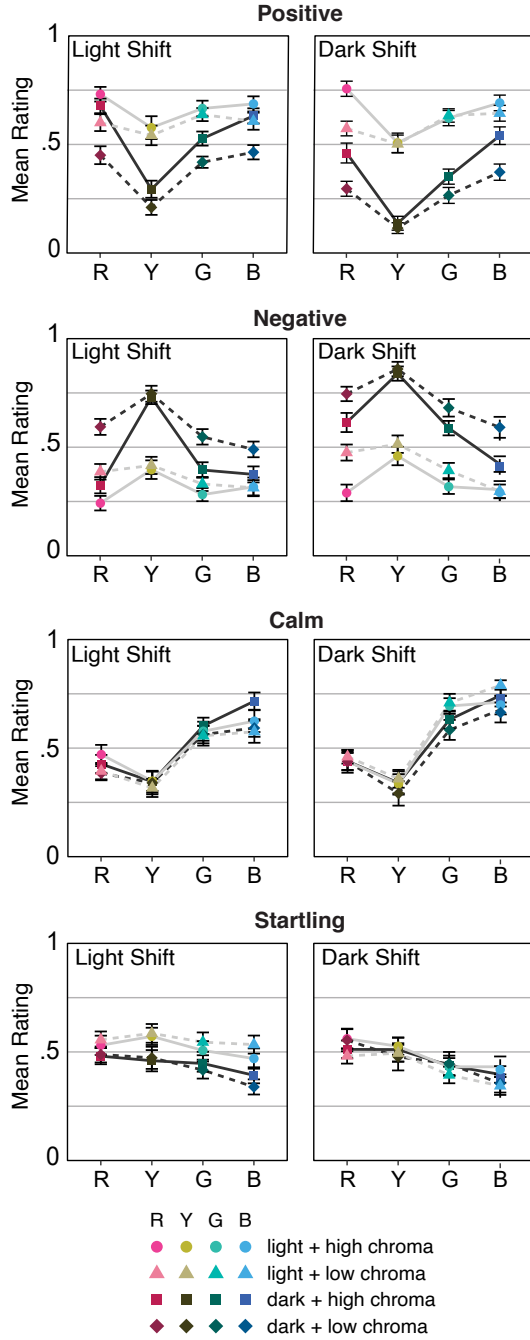


Fig. 9: Mean color-concept association ratings for light shifted (left) and dark shifted (right) colormaps from Exp. 2. Data are plotted in the same manner as Fig. 5 from Exp. 1.

Also as in Exp. 1, *negative* was associated with darker, lower chroma, yellower colormaps, and increasing chroma especially reduced negativity when colors were bluer and darker. Unlike Exp. 1, an effect of red/green and a red/green $\times$ chroma interaction indicated redder colors were more negative, particularly when chroma was low. Again, although these effects were not significant in Exp. 1, the weights were in the same direction (Table S.4).

Finally, like Exp. 1, *calm* was associated with bluer colormaps, but unlike Exp. 1 it was associated with greener, higher chroma colormaps. The weight for red/green was in the same direction as Exp. 1 but not for chroma (Table S.4), and the reason is unclear. Moreover, in Exp. 2 chroma interacted with lightness, red/green, blue/yellow, but these interactions are difficult to interpret as the curves for *calm* largely overlap in Fig. 9. New to Exp. 2, *startling* was associated with yellower colormaps, suggesting it is somewhat opposite of *calm*.

Overall, the mean associations between matched color scales and emotions across Exp. 1 and 2 (averaged over shift condition) were strongly correlated (*positive*: $r = 0.97, p < 0.001$, *negative*: $r = 0.98, p < 0.001$, and *calm*: $r = 0.90, p < 0.001$). The remaining analyses separated the data by shift to test the data-dependence hypothesis.

Effects of shift. Overall, light shifted colormaps were more associated with *positive* and dark shifted colormaps were more associated with *negative* (Figs. 6B and 9). Shift $\times$ lightness interactions for both concepts indicated that the effect of color scale lightness was greater for dark shifted maps (i.e., dark shifted colormaps constructed from darker color scales were especially less *positive* and more *negative*). Dark shifting also amplified the extent to which yellower colormaps were more *negative* than bluer colormaps (shift $\times$ yellow/blue interaction), especially for darker, higher chroma colors (shift $\times$ lightness $\times$ chroma $\times$ yellow/blue interaction). *Negative* also had a 4-way interaction with red/green, such that dark shifting magnified differences due to lightness and chroma for redder colormaps, which were less extreme for greener colormaps (Fig. 9). Together, these results contribute to the narrative that effects of color dimensions on *positive/negative* associations are generally more extreme when colors are darker, as determined by darker colors scales or dark shifting the underlying dataset.

Turning to the other two emotions, dark shifting overall increased the associations with *calm*, especially for colormaps that were bluer (shift $\times$ yellow/blue interaction). Fig. 6B shows additional higher order interactions for *calm*, but they are difficult to interpret because of the overlapping pattern of associations (Fig. 9). For *startling*, shift had no overall effect, but it interacted with lightness—increasing lightness made colormaps especially more startling when they were light shifted.

The effects of shift reported in this section support the data-dependence hypothesis, suggesting that modeling affective connotation of colormaps based only on colors in the color scale is insufficient. It is necessary to account for the frequency colors appear in the colormap as determined by the underlying data. To test this implication directly, we compared two mixed-effect regression models for each emotion. The ‘data-agnostic’ models predicted association ratings using the mean L^* , C^* , a^* , and b^* coordinates of the color scales used to construct the colormaps (same as the models in Exp. 1). The ‘data-aware’ models used the average L^* , C^* , a^* , and b^* coordinates of the pixels in each colormap, which weighted the colors in the color scales by the frequency they appeared in the colormaps. To enable model comparisons, we fit these models using maximum likelihood (ML) as opposed to restricted maximum likelihood (REML) as the objective function. Table 1 shows that the data-aware models outperformed the data-agnostic models for all four emotions according to AIC ($\downarrow$ better), BIC ($\downarrow$ better), and conditional R^2 ($\uparrow$ better) metrics. Individual model results are summarized in Tables S.7 and S.8

Table 1: Model comparisons between mixed effects models using color scale and colormap predictors ($df = 18$ for all models).

Emotion	Model Type	AIC	BIC	Cond. R^2
Positive	Data-agnostic	-441.05	-338.72	0.518
	Data-aware	-566.66	-464.33	0.558
Negative	Data-agnostic	-421.51	-319.17	0.538
	Data-aware	-628.47	-526.14	0.593
Calm	Data-agnostic	-130.48	-28.14	0.530
	Data-aware	-361.37	-259.03	0.593
Startling	Data-agnostic	-305.99	-203.66	0.489
	Data-aware	-395.13	-296.60	0.534

4.2.2 Analysis of underlying dimensional structure

Fig. 7B shows the results of a PCA on the mean association ratings from Exp. 2. Each of the 32 points in this space represents the mean of all colormaps constructed from each color scale $\times$ shift level. Light shifted and dark shifted colormaps produced using the same color scale are connected by a line colored according to the middle color of the color scale. If shift had no effect, then colormaps produced using the same color scale would share the same position in the space. However, that was not the case, and some colormaps even traversed different quadrants of the space. For example, colormaps produced from the

dark, high chroma red color scale (Fig. 8A) were positive when light shifted and became negative when dark shifted (Fig. 7B).

The dimensions of this affective space were well mapped to valence (positive vs. negative) and arousal/activation (activated vs. deactivated) as in Exp. 1; see Supplementary Material Fig. S.6 for correlations. Using multiple linear regression, we characterized the color variation along each PC (see Table S.5), but instead of data-agnostic predictors we used data-aware predictors, shown to fit the data better (Table 1). The PCs had similar structure as in Exp 1—PC1 (valence) increased with lightness, chroma, and blueness (adjusted model $R^2 = 0.716$) and PC2 (activation) increased with lightness, redness, and yellowness (adjusted model $R^2 = 0.832$). The one difference is valence was not related to red/green in Exp. 1 but increased with greenness in Exp. 2.

Summary. Supporting the data-dependence hypothesis, affective connotation depended not only on the color scales used to construct colormaps, but also the proportion that the colors appear in the colormap, as determined by the underlying dataset. The colors that were more prevalent contributed more to the affective connotation of colormaps. These results imply that for a designer to anticipate the affective connotation of colormaps, it is important that they use a data-aware approach that accounts for how the colors they choose will appear in the context of the visualizations, given the relative frequency of values in the data.

5 GENERAL DISCUSSION

Our study was motivated by the premises that (a) affective connotations of colors are largely driven by lightness level, and (b) for colormap data visualizations to reveal fine details in spatial structure, they need to vary in lightness. Our primary aim was to investigate whether colormaps that vary in lightness to support spatial vision can communicate distinct affective connotations. Indeed, we found systematic color-emotion associations, similar to previous findings for single colors [15, 16, 26, 68, 84]. For example, positive emotions were associated with lighter, higher chroma, and bluer colormaps, whereas negative emotions had the opposite associations. And, similar to previous results for color palettes [6], cooler colormaps were associated with *calm*.

However, we also found differences from previous work on color palettes. For example, whereas yellower color palettes were *positive*, *playful*, and *exciting* [6], we found that those emotions were associated with bluer colormaps. This yellow/blue discrepancy can be understood in terms of the role of chroma and lightness. In Bartram et al.'s *exciting*, *positive*, and *playful* color palettes, the warm colors were all relatively light and high in chroma, but for the present colormaps, the colors varied widely in lightness and chroma (Fig. 1). As yellower hues darken they can no longer be high in chroma [45, 85], and a decrease in both lightness and chroma makes them especially negative [28, 68]. In contrast, as bluer hues darken, they can become higher in chroma [45, 85], and increasing chroma as blues darken helps preserve the positive association [68]. It follows that for colormaps varying in lightness, bluer colormaps should be more positive than yellower colormaps because they shift less toward negative as the colors get darker. These results imply that a polychromatic color scale like *viridis*, which has light, high chroma yellows and dark, high chroma blues, would be strongly positive. Indeed, evidence suggests colormaps produced with *viridis* evoke positive emotional responses [89].

Our study also aimed to test the data-dependence hypothesis, proposing that affective connotation depends not only *which* colors are present in the visualization, but also *how much* those colors appeared in the image as determined by the underlying dataset. Supporting this hypothesis, Exp. 2 showed that shifting the underlying dataset modulated color-emotion associations (e.g., light shifted colormaps were associated more with *positive* and less with *negative* or *calm* than dark shifted maps). Moreover, the pattern of associations was better predicted by data-aware models that accounted for the color distributions of pixels within the colormap images compared to data-agnostic models that gave equal weight to every color in the color scale.

In the present study, shifting the underlying dataset primarily modulated the lightness of pixels in the maps because the color scales were roughly monochromatic and most of them varied in chroma in an inverted U function (the high chroma colors were in the middle of the

scale; Fig. S.1). However, if a polychromatic color scale was used (e.g., *viridis*), shift would influence hue representation as well as lightness. Moreover, data can have all kinds of distributions that influence which colors are represented in the colormaps. For example, if we had shifted the data to the mid-range, the colors would have had mid-level lightness and would have been higher in chroma for most color scales in our set.

These points emphasize the importance of data-aware visualization design, which accounts for not only which design features encode data (e.g., colors, shapes, textures), but also how those design features are instantiated in a visualization, given the numeric and semantic properties of the underlying dataset. A data-aware approach for designing colormaps to have particular affective connotations might look like this: (a) identify colors of the desired affective connotation, (b) examine the histogram of the dataset, and (c) apply a color scale such that the colors identified in (a) get mapped to the most frequent colors within the dataset. Other examples of data-aware colormap design involve (a) selecting/modifying color scales so they help observers perform specific kinds of tasks or highlight what Zeng et al. [88] call 'meaningful features' of data [59, 63, 79, 88] and (b) examining the spatial structure of data [71, 91] and the concepts that the data represent [69, 73] to anticipate whether observers will infer that darker or lighter colors represent larger quantities in the colormap. Here, our emphasis is on colormaps, but questions remain concerning the importance of data-awareness for categorical visualizations. Blair et al. [9] suggested links between color and emotion for categorical visualizations are similar to those of isolated color palettes, implying data agnosticity. But, data-awareness may become more important for categorical visualizations if there is greater discrepancy in the area occupied by each color (e.g., different sized bars in bar charts or regions in choropleth maps).

Our study raises additional open questions for future work. First, how do color and spatial factors combine to influence affective connotations of colormaps? Here, we controlled the affective connotation of spatial structure by choosing maps that had the most similar affective connotation in grayscale. Future work can test whether spatial and color factors influence affective connotation independently or interact in interesting ways. A second question concerns potential implications of our results for judgment and decision-making pertaining to visualizations. Anderson and Robinson [2] reported that congruence between visualization topic and colors increased judgments of appropriateness. Do such design factors influence down-stream decisions (e.g., likelihood to support a cause) or the interpretations of data in storytelling contexts [10, 38]? Finally, our study was conducted in the US among college undergraduates, but how might these results vary across cultures? Although studies have found cross-cultural commonalities in color-emotion associations, there are also systematic cultural differences [1, 16, 25, 28, 30, 47, 50, 77, 86]. For example, Xin et al. [86] found that color-emotion associations for participants from Thailand were especially driven by lightness, compared to those from Hong Kong and Japan. A comprehensive understanding of affective connotation of color in visualizations will require a cross-cultural approach.

6 CONCLUSION

We found that colormap visualizations that vary in lightness to support fine-detail spatial vision can communicate clear affective connotations. Affective connotation depends not only on which colors appear in a colormap but also on color frequencies as determined by the underlying dataset, emphasizing the importance of data-aware design. This study was a step toward understanding affective connotation in colormap design, and future work can extend this approach by studying data-aware visualization design involving polychromatic color scales, interactions between color and spatial factors, and variations across populations.

ACKNOWLEDGMENTS

We thank Danielle Szafir, Psyche Loui, Domicile Jonauskaitė, and Melissa Schoenlein. This work was supported in part by NSF BCS-1945303. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. Use of AI: Github Copilot helped write code for experiment tasks and Claude helped format tables in LaTeX.

REFERENCES

- [1] F. M. Adams and C. E. Osgood. A cross-cultural study of the affective meanings of color. *J. Cross-Cult. Psychol.*, 4(2):135–156, 1973. doi: [10.1177/002202217300400201](https://doi.org/10.1177/002202217300400201) 6, 9
- [2] C. L. Anderson and A. C. Robinson. Affective congruence in visualization design: Influences on reading categorical maps. *IEEE Trans. Vis. Comput. Graph.*, 28(8):2867–2878, 2022. doi: [10.1109/tvcg.2021.3050118](https://doi.org/10.1109/tvcg.2021.3050118) 1, 3, 9
- [3] J. R. Antes and K.-T. Chang. An empirical analysis of the design principles for quantitative and qualitative area symbols. *Cartogr. Geogr. Inf. Sys.*, 17(4):271–277, 1990. doi: [10.1559/152304090783805663](https://doi.org/10.1559/152304090783805663) 1
- [4] R. H. Baayen, D. J. Davidson, and D. M. Bates. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.*, 59(4):390–412, 2008. doi: [10.1016/j.jml.2007.12.005](https://doi.org/10.1016/j.jml.2007.12.005) 5
- [5] L. F. Barrett and C. Westlin. Navigating the science of emotion. In H. L. Meiselman, ed., *Emotion Measurement*, pp. 39–84. Elsevier, 2021. doi: [10.1016/b978-0-12-821124-3.00002-8](https://doi.org/10.1016/b978-0-12-821124-3.00002-8) 2
- [6] L. Bartram, A. Patra, and M. Stone. Affective color in visualization. In *Proc. CHI*, pp. 1364–1374, 2017. doi: [10.1145/3025453.3026041](https://doi.org/10.1145/3025453.3026041) 1, 2, 3, 6, 7, 9
- [7] D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, 67(1):1–48, 2015. doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01) 5
- [8] A. M. Belfi and P. Loui. Musical anhedonia and rewards of music listening: Current advances and a proposed model. *Ann. N. Y. Acad. Sci.*, 1464(1):99–114, 2020. doi: [10.1024/1016-264x/a000292](https://doi.org/10.1024/1016-264x/a000292) 2
- [9] C. Blair, X. Wang, and C. Perin. Quantifying emotional responses to immutable data characteristics and designer choices in data visualizations. *IEEE Trans. Vis. Comput. Graph.*, 2025. doi: [10.1109/tvcg.2024.3456361](https://doi.org/10.1109/tvcg.2024.3456361) 1, 3, 9
- [10] J. Boy, A. V. Pandey, J. Emerson, M. Satterthwaite, O. Nov, and E. Bertini. Showing people behind data: Does anthropomorphizing visualizations elicit more empathy for human rights data? In *Proc. CHI*, pp. 5462–5474, 2017. doi: [10.1145/3025453.3025512](https://doi.org/10.1145/3025453.3025512) 9
- [11] C. A. Brewer. Color use guidelines for mapping and visualization. In A. M. MacEachren and D. R. F. Taylor, eds., *Visualization in Modern Cartography*, pp. 123–148. Elsevier Science Inc., Tarrytown, 1994. doi: [10.1016/b978-0-08-042415-6.50014-4](https://doi.org/10.1016/b978-0-08-042415-6.50014-4) 1
- [12] R. Bujack, T. L. Turton, F. Samsel, C. Ware, D. H. Rogers, and J. Ahrens. The good, the bad, and the ugly: A theoretical framework for the assessment of continuous colormaps. *IEEE Trans. Vis. Comput. Graph.*, 24(1):923–933, 2018. doi: [10.1109/tvcg.2017.2743978](https://doi.org/10.1109/tvcg.2017.2743978) 1
- [13] G. L. Collier. Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motiv. Emot.*, 20(1):1–32, 1996. doi: [10.1007/bf02251005](https://doi.org/10.1007/bf02251005) 1
- [14] D. J. Cuff. Colour on temperature maps. *Cartogr. J.*, 10(1):17–21, 1973. doi: [10.1179/caj.1973.10.1.17](https://doi.org/10.1179/caj.1973.10.1.17) 1
- [15] N. Dael, M.-N. Perseguers, C. Marchand, J.-P. Antonietti, and C. Mohr. Put on that colour, it fits your emotion: Colour appropriateness as a function of expressed emotion. *QJ Exp. Psychol.*, 69(8):1619–1630, 2016. doi: [10.1080/17470218.2015.1090462](https://doi.org/10.1080/17470218.2015.1090462) 1, 2, 9
- [16] R. D’Andrade and M. Egan. The colors of emotion. *Am. Ethnol.*, 1(1):49–63, 1974. doi: [10.1525/ae.1974.1.1.02a00030](https://doi.org/10.1525/ae.1974.1.1.02a00030) 1, 2, 3, 9
- [17] J. R. De Leeuw. jspsych: A javascript library for creating behavioral experiments in a web browser. *Behav. Res. Methods*, 47:1–12, 2015. doi: [10.1201/b19402-15](https://doi.org/10.1201/b19402-15) 4
- [18] J. R. de Leeuw, R. A. Gilbert, and B. Luchterhandt. jspsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *J. Open Source Softw.*, 8(85):5351, 2023. doi: [10.21105/joss.05351](https://doi.org/10.21105/joss.05351) 4
- [19] K. K. De Valois and F. Kooi. The role of color in spatial vision. In L. Harris and M. Jenkin, eds., *Spatial Vision in Humans and Robots*, pp. 149–159, 1993. 2
- [20] R. L. DeValois and K. K. DeValois. *Spatial Vision*. Oxford university press, 1990. 2
- [21] C. C. Gramazio, D. H. Laidlaw, and K. B. Schloss. Colorgorical: Creating discriminable and preferable color palettes for information visualization. *IEEE Trans. Vis. Comput. Graph.*, 23(1):521–530, 2017. doi: [10.1109/tvcg.2016.2598918](https://doi.org/10.1109/tvcg.2016.2598918) 4
- [22] G. Grand, I. A. Blank, F. Pereira, and E. Fedorenko. Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nat. Hum. Behav.*, 6(7):975–987, 2022. doi: [10.1038/s41562-022-01316-8](https://doi.org/10.1038/s41562-022-01316-8) 12
- [23] N. L. Harris, D. A. Gibbs, A. Baccini, R. A. Birdsey, S. De Bruin, M. Farina, L. Fatoyinbo, M. C. Hansen, M. Herold, R. A. Houghton, et al. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Change*, 11(3):234–240, 2021. doi: [10.1038/s41558-020-00976-6](https://doi.org/10.1038/s41558-020-00976-6) 1, 2, 4
- [24] M. Harrower and C. A. Brewer. ColorBrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, 2003. doi: [10.1559/152304003100011126](https://doi.org/10.1559/152304003100011126) 2
- [25] D. Jonauskaitė, A. Abu-Akel, N. Dael, D. Oberfeld, A. M. Abdel-Khalek, A. S. Al-Rasheed, J.-P. Antonietti, V. Bogushevskaya, A. Chamseddine, E. Chkonia, et al. Universal patterns in color-emotion associations are further shaped by linguistic and geographic proximity. *Psychol. Sci.*, 31(10):1245–1260, 2020. doi: [10.1177/0956797620948810](https://doi.org/10.1177/0956797620948810) 9
- [26] D. Jonauskaitė, B. Althaus, N. Dael, E. Dan-Glauser, and C. Mohr. What color do you feel? Color choices are driven by mood. *Color Res. Appl.*, 44(2):272–284, 2019. doi: [10.1002/col.22327](https://doi.org/10.1002/col.22327) 1, 2, 9
- [27] D. Jonauskaitė, L. Camenzind, C. A. Parraga, C. N. Diouf, M. M. Ducommun, L. Müller, M. Norberg, and C. Mohr. Colour-emotion associations in individuals with red-green colour blindness. *PeerJ*, 9:e11180, 2021. doi: [10.7717/peerj.11180](https://doi.org/10.7717/peerj.11180) 3
- [28] D. Jonauskaitė and C. Mohr. Do we feel colours? A systematic review of 128 years of psychological research linking colours and emotions. *Psychon. Bull. Rev.*, pp. 1–30, 2025. doi: [10.3758/s13423-024-02615-z](https://doi.org/10.3758/s13423-024-02615-z) 1, 2, 3, 6, 9
- [29] D. Jonauskaitė, G. F. M. Spagnulo, D. Epicoco, and C. Mohr. Beware! colour and emotion correspondences are rarely about feelings. In C. P. Biggam, D. Jonauskaitė, D. Mylonas, and M. Uuskuula, eds., *Progress in Colour Studies*. John Benjamins Publishing Company, in press. 2
- [30] D. Jonauskaitė, J. Wicker, C. Mohr, N. Dael, J. Havelka, M. Papadatou-Pastou, M. Zhang, and D. Oberfeld. A machine learning approach to quantify the specificity of colour–emotion associations and their cultural differences. *R. Soc. Open Sci.*, 6(9):190741, 2019. doi: [10.1098/rsos.190741](https://doi.org/10.1098/rsos.190741) 9
- [31] N. Kathios, A. D. Patel, and P. Loui. Musical anhedonia, timbre, and the rewards of music listening. *Cognition*, 243:105672, 2024. doi: [10.1016/j.cognition.2023.105672](https://doi.org/10.1016/j.cognition.2023.105672) 2
- [32] N. Kaya and H. Epps. Relationship between color and emotion: A study of college students. *J. Coll. Stud.*, 38(3):396, 2004. 1
- [33] G. Kindlmann, E. Reinhard, and S. Creem. Face-based luminance matching for perceptual colormap generation. In *IEEE Visualization, 2002.*, pp. 299–306, 2002. doi: [10.1109/VISUAL.2002.1183788](https://doi.org/10.1109/VISUAL.2002.1183788) 2
- [34] X. Lan, Y. Shi, Y. Wu, X. Jiao, and N. Cao. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Trans. Vis. Comput. Graph.*, 28(1):933–943, 2021. doi: [10.1109/TVCG.2021.3114775](https://doi.org/10.1109/TVCG.2021.3114775) 1
- [35] X. Lan, Y. Shi, Y. Zhang, and N. Cao. Smile or scowl? Looking at infographic design through the affective lens. *IEEE Trans. Vis. Comput. Graph.*, 27(6):2796–2807, 2021. doi: [10.1109/tvcg.2021.3074582](https://doi.org/10.1109/tvcg.2021.3074582) 1
- [36] X. Lan, Y. Wu, and N. Cao. Affective visualization design: Leveraging the emotional impact of data. *IEEE Trans. Vis. Comput. Graph.*, 30(1):1–11, 2024. doi: [10.1109/tvcg.2023.3327385](https://doi.org/10.1109/tvcg.2023.3327385) 1, 2
- [37] X. Lan, Y. Wu, Q. Chen, and N. Cao. The chart excites me! exploring how data visualization design influences affective arousal. *arXiv preprint arXiv:2211.03296*, 2022. doi: [10.48550/arXiv.2211.03296](https://doi.org/10.48550/arXiv.2211.03296) 1
- [38] J. Liem, C. Perin, and J. Wood. Structure and empathy in visual data storytelling: Evaluating their influence on attitude. *Computer Graphics Forum*, 39(3):277–289, 2020. doi: [10.1111/cgf.13980](https://doi.org/10.1111/cgf.13980) 9
- [39] S. Lin, J. Fortuna, C. Kulkarni, M. Stone, and J. Heer. Selecting semantically-resonant colors for data visualization. *Computer Graphics Forum*, 32:401–410, 2013. doi: [10.1111/cgf.12127](https://doi.org/10.1111/cgf.12127) 3
- [40] M. S. Livingstone and D. H. Hubel. Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J. Neurosci.*, 7(11):3416–3468, 1987. doi: [10.1523/jneurosci.07-11-03416.1987](https://doi.org/10.1523/jneurosci.07-11-03416.1987) 2
- [41] P. Loui, S. Patterson, M. E. Sachs, Y. Leung, T. Zeng, and E. Przsinda. White matter correlates of musical anhedonia: implications for evolution of music. *Front. Psychol.*, 8:1664, 2017. doi: [10.3389/fpsyg.2017.01664](https://doi.org/10.3389/fpsyg.2017.01664) 2
- [42] M. McGranaghan. Ordering choropleth map symbols: The effect of background. *Am. Cartog.*, 16(4):279–285, 1989. doi: [10.1559/152304089783813918](https://doi.org/10.1559/152304089783813918) 1
- [43] K. Moreland. Diverging color maps for scientific visualization. In *International Symposium on Visual Computing*, pp. 92–103. Springer, 2009. doi: [10.1007/978-3-642-10520-3_9](https://doi.org/10.1007/978-3-642-10520-3_9) 1
- [44] K. Mukherjee, B. Yin, B. E. Sherman, L. Lessard, and K. B. Schloss. Context matters: A theory of semantic discriminability for perceptual encoding systems. *IEEE Trans. Vis. Comput. Graph.*, 28(1):697–706,

2022. doi: [10.1109/tvcg.2021.3114780](https://doi.org/10.1109/tvcg.2021.3114780) 3, 4
- [45] A. Munsell. *A Grammar of Color*. New York: Van Nostrand Reinhold, 1921. doi: [10.7312/merc90424-002](https://doi.org/10.7312/merc90424-002) 9
- [46] P. Nardini, M. Chen, F. Samsel, R. Bujack, M. Böttinger, and G. Scheuermann. The making of continuous colormaps. *IEEE Trans. Vis. Comput. Graph.*, 27(6):3048–3063, 2021. doi: [10.1109/tvcg.2019.2961674](https://doi.org/10.1109/tvcg.2019.2961674) 1
- [47] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Res. Appl.*, 29(3):232–240, 2004. doi: [10.1002/col.20047](https://doi.org/10.1002/col.20047) 9
- [48] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. Part II: colour emotions for two-colour combinations. *Color Res. Appl.*, 29(4):292–298, 2004. doi: [10.1002/col.20024](https://doi.org/10.1002/col.20024) 1
- [49] L.-C. Ou, M. Ronnier Luo, P.-L. Sun, N.-C. Hu, H.-S. Chen, S.-S. Guan, A. Woodcock, J. L. Caivano, R. Huertas, A. Treméau, et al. A cross-cultural comparison of colour emotion for two-colour combinations. *Color Res. Appl.*, 37(1):23–43, 2012. doi: [10.1002/col.20648](https://doi.org/10.1002/col.20648) 1
- [50] S. E. Palmer, K. B. Schloss, Z. Xu, and L. R. Prado-León. Music–color associations are mediated by emotion. *Proc. Natl. Acad. Sci.*, 110(22):8836–8841, 2013. doi: [10.1073/pnas.1212562110](https://doi.org/10.1073/pnas.1212562110) 5, 6, 9
- [51] A. D. Pazda, C. A. Thorstenson, and A. K. Fetterman. Colorfulness influences perceptions of valence and arousal. *J. Exp. Psychol. Gen.*, 153(1):145, 2024. doi: [10.1037/xge0001484](https://doi.org/10.1037/xge0001484) 1
- [52] J. Pennington, R. Socher, and C. D. Manning. GloVe: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014. doi: [10.3115/v1/d14-1162](https://doi.org/10.3115/v1/d14-1162) 7, 12
- [53] F. Pereira, B. Lou, B. Pritchett, S. Ritter, S. J. Gershman, N. Kanwisher, M. Botvinick, and E. Fedorenko. Toward a universal decoder of linguistic meaning from brain activation. *Nat. Commun.*, 9(1):963, 2018. doi: [10.1038/s41467-018-03068-4](https://doi.org/10.1038/s41467-018-03068-4) 12
- [54] J. Posner, J. A. Russell, and B. S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.*, 17(3):715–734, 2005. doi: [10.1017/s0954579405050340](https://doi.org/10.1017/s0954579405050340) 3, 6, 7
- [55] R. R Core Team et al. R: A language and environment for statistical computing, 2013. 5
- [56] K. Reda, P. Nalawade, and K. Ansah-Koi. Graphical perception of continuous quantitative maps: the effects of spatial frequency and colormap design. In *Proc. CHI*, p. 272. ACM, 2018. doi: [10.1145/3173574.3173846](https://doi.org/10.1145/3173574.3173846) 1
- [57] K. Reda, A. A. Salvi, J. Gray, and M. E. Papka. Color nameability predicts inference accuracy in spatial visualizations. *Computer Graphics Forum*, 40:49–60, 2021. doi: [10.1111/cgf.14288](https://doi.org/10.1111/cgf.14288) 1
- [58] K. Reda and D. A. Szafrir. Rainbows revisited: modeling effective colormap design for graphical inference. *IEEE Trans. Vis. Comput. Graph.*, 27(2):1032–1042, 2020. doi: [10.1109/TVCG.2020.3030439](https://doi.org/10.1109/TVCG.2020.3030439) 1
- [59] B. E. Rogowitz and L. A. Treinish. How not to lie with visualization. *Comput. Phys.*, 10(3):268–273, 1996. doi: [10.1063/1.4822401](https://doi.org/10.1063/1.4822401) 1, 2, 9
- [60] B. E. Rogowitz and L. A. Treinish. Data visualization: the end of the rainbow. *IEEE Spectrum*, 35(12):52–59, 1998. doi: [10.1109/6.736450](https://doi.org/10.1109/6.736450) 1
- [61] J. A. Russell. A circumplex model of affect. *J. Pers. Soc. Psychol.*, 39(6):1161, 1980. doi: [10.1037/h0077714](https://doi.org/10.1037/h0077714) 3
- [62] F. Samsel, L. Bartram, and A. Bares. Art, affect and color: Creating engaging expressive scientific visualization. In *2018 IEEE VIS Arts Program (VISAP)*, pp. 1–9. IEEE, 2018. doi: [10.1109/visap45312.2018.9046053](https://doi.org/10.1109/visap45312.2018.9046053) 1, 3
- [63] F. Samsel, S. Klaassen, and D. H. Rogers. ColorMoves: Real-time interactive colormap construction for scientific visualization. *IEEE Comput. Graph. Appl.*, 38(1):20–29, 2018. doi: [10.1109/mcg.2018.011461525](https://doi.org/10.1109/mcg.2018.011461525) 9
- [64] M. Satoh, T. Nakase, K. Nagata, and H. Tomimoto. Musical anhedonia: selective loss of emotional experience in listening to music. *Neurocase*, 17(5):410–417, 2011. doi: [10.1080/13554794.2010.532139](https://doi.org/10.1080/13554794.2010.532139) 2
- [65] K. B. Schloss, C. C. Gramazio, A. T. Silverman, M. L. Parker, and A. S. Wang. Mapping color to meaning in colormap data visualizations. *IEEE Trans. Vis. Comput. Graph.*, 25(1):810–819, 2019. doi: [10.1109/tvcg.2018.2865147](https://doi.org/10.1109/tvcg.2018.2865147) 1
- [66] K. B. Schloss, Z. Leggon, and L. Lessard. Semantic discriminability for visual communication. *IEEE Trans. Vis. Comput. Graph.*, 27(2):1022–1031, 2021. doi: [10.1109/tvcg.2020.3030434](https://doi.org/10.1109/tvcg.2020.3030434) 3
- [67] K. B. Schloss, L. Lessard, C. S. Walmsley, and K. Foley. Color inference in visual communication: the meaning of colors in recycling. *Cogn. Res. Princ. Implic.*, 3(1):5, 2018. doi: [10.1186/s41235-018-0090-y](https://doi.org/10.1186/s41235-018-0090-y) 3
- [68] K. B. Schloss, C. Witzel, and L. Y. Lai. Blue hues don't bring the blues: questioning conventional notions of color–emotion associations. *J. Opt. Soc. Am. A*, 37(5):813–824, 2020. doi: [10.1364/josaa.383588](https://doi.org/10.1364/josaa.383588) 1, 2, 3, 5, 7, 9
- [69] M. A. Schoenlein, J. Campos, K. J. Lande, L. Lessard, and K. B. Schloss. Unifying effects of direct and relational associations for visual communication. *IEEE Trans. Vis. Comput. Graph.*, 29(1):385–395, 2023. doi: [10.1109/TVCG.2022.3209443](https://doi.org/10.1109/TVCG.2022.3209443) 1, 9
- [70] S. Schömbis, J. Pan, Y. Zhang, J. Goncalves, and W. Johal. Facevis: Exploring a robot's face for affective visualisation design. In *Proc. CHI*, pp. 1–10, 2024. doi: [10.1145/3613905.3650910](https://doi.org/10.1145/3613905.3650910) 1
- [71] S. C. Sibrel, R. Rathore, L. Lessard, and K. B. Schloss. The relation between color and spatial structure for interpreting colormap data visualizations. *J. Vis.*, 20(12):7–7, 2020. doi: [10.1167/jov.20.12.7](https://doi.org/10.1167/jov.20.12.7) 1, 9
- [72] S. Smart, K. Wu, and D. A. Szafrir. Color crafting: Automating the construction of designer quality color ramps. *IEEE Trans. Vis. Comput. Graph.*, 26(1):1215–1225, 2020. doi: [10.1109/tvcg.2019.2934284](https://doi.org/10.1109/tvcg.2019.2934284) 1, 2, 3, 4
- [73] A. Soto, M. A. Schoenlein, and K. B. Schloss. More of what? Dissociating effects of conceptual and numeric mappings on interpreting colormap data visualizations. *Cogn. Res. Princ. Implic.*, 8(1):38, 2023. doi: [10.1186/s41235-023-00482-1](https://doi.org/10.1186/s41235-023-00482-1) 1, 9
- [74] M. Stone, D. A. Szafrir, and V. Setlur. An engineering model for color difference as a function of size. In *Color and Imaging Conference*, vol. 2014, pp. 253–258. Society for Imaging Science and Technology, 2014. doi: [10.2352/cic.2014.22.1.art00045](https://doi.org/10.2352/cic.2014.22.1.art00045) 4
- [75] H. J. Suk and H. Intel. Emotional response to color across media. *Color Res. Appl.*, 35(1):64–77, 2010. doi: [10.1002/col.20554](https://doi.org/10.1002/col.20554) 1
- [76] D. A. Szafrir. Modeling color difference for visualization design. *IEEE Trans. Vis. Comput. Graph.*, 24(1):392–401, 2018. doi: [10.1109/tvcg.2017.2744359](https://doi.org/10.1109/tvcg.2017.2744359) 4
- [77] D. S. Y. Tham, P. T. Sowden, A. Grandison, A. Franklin, A. K. W. Lee, M. Ng, J. Park, W. Pang, and J. Zhao. A systematic investigation of conceptual color associations. *J. Exp. Psychol. Gen.*, 149(7):1311, 2020. doi: [10.1037/xge0000703](https://doi.org/10.1037/xge0000703) 9
- [78] C. A. Thorstenson, A. J. Elliot, A. D. Pazda, D. I. Perrett, and D. Xiao. Emotion-color associations in the context of the face. *Emotion*, 18(7):1032–1042, 2018. doi: [10.1037/emo0000358](https://doi.org/10.1037/emo0000358) 1
- [79] C. Tominski, G. Fuchs, and H. Schumann. Task-driven color coding. In *2008 12th International Conference Information Visualisation*, pp. 373–380. IEEE, 2008. doi: [10.1109/iv.2008.24](https://doi.org/10.1109/iv.2008.24) 9
- [80] P. Valdez and A. Mehrabian. Effects of color on emotions. *J. Exp. Psychol. Gen.*, 123(4):394, 1994. doi: [10.1037//0096-3445.123.4.394](https://doi.org/10.1037//0096-3445.123.4.394) 1, 6
- [81] C. Ware. Color sequences for univariate maps: Theory, experiments and principles. *IEEE Comput. Graph. Appl.*, 8(5):41–49, 1988. doi: [10.1109/38.7760](https://doi.org/10.1109/38.7760) 2
- [82] L. B. Wexner. The degree to which colors (hues) are associated with mood-tones. *J. Appl. Psychol.*, 38(6):432, 1954. doi: [10.1037/h0062181](https://doi.org/10.1037/h0062181) 1
- [83] L. Wilms and D. Oberfeld. Color and emotion: effects of hue, saturation, and brightness. *Psychol. Res.*, 82(5):896–914, 2018. doi: [10.1007/s00426-017-0880-8](https://doi.org/10.1007/s00426-017-0880-8) 1
- [84] B. Wright and L. Rainwater. The meanings of color. *J. Gen. Psychol.*, 67(1):89–99, 1962. doi: [10.1080/00221309.1962.9711531](https://doi.org/10.1080/00221309.1962.9711531) 1, 2, 3, 9
- [85] G. Wyszecki and W. S. Stiles. *Color Science*, vol. 8. Wiley New York, 1982. doi: [10.1002/col.5080070104](https://doi.org/10.1002/col.5080070104) 9
- [86] J. H. Xin, K. M. Cheng, G. Taylor, T. Sato, and A. Hansuebsai. Cross-regional comparison of colour emotions Part I: Quantitative analysis. *Color Res. Appl.*, 29(6):451–457, 2004. doi: [10.1002/col.20062](https://doi.org/10.1002/col.20062) 9
- [87] S. Zeller, F. Samsel, and L. Bartram. Affective, hand-sculpted glyph forms for engaging and expressive scientific visualization. In *IEEE VIS Arts Program (VISAP)*, pp. 127–136. IEEE, 2022. doi: [10.1109/visap57411.2022.00025](https://doi.org/10.1109/visap57411.2022.00025) 1
- [88] Q. Zeng, Y. Zhao, Y. Wang, J. Zhang, Y. Cao, C. Tu, I. Viola, and Y. Wang. Data-driven colormap adjustment for exploring spatial variations in scalar fields. *IEEE Trans. Vis. Comput. Graph.*, 28(12):4902–4917, 2021. doi: [10.1109/TVCG.2021.3109014](https://doi.org/10.1109/TVCG.2021.3109014) 1, 9
- [89] Y. Zhang, M. Fjeld, M. Fratacangeli, A. Said, and S. Zhao. Affective colormap design for accurate visual comprehension in industrial tomography. *Sensors*, 21(14):4766, 2021. doi: [10.3390/s21144766](https://doi.org/10.3390/s21144766) 1, 3, 9
- [90] L. Zhou and C. D. Hansen. A survey of colormaps in visualization. *IEEE Trans. Vis. Comput. Graph.*, 22(8):2051–2069, 2016. doi: [10.1109/tvcg.2015.2489649](https://doi.org/10.1109/tvcg.2015.2489649) 1
- [91] C. Zimnicki, C. Tseng, D. A. Szafrir, and K. B. Schloss. Effects of data distribution and granularity on color semantics for colormap data visualizations. In *2023 IEEE Visualization and Visual Analytics (VIS)*, pp. 11–15. IEEE, 2023. doi: [10.1109/vis54172.2023.00011](https://doi.org/10.1109/vis54172.2023.00011) 9

S SUPPLEMENTARY MATERIAL

S.1 Participants in pilot study on finding gray maps with similar affective connotation

The participants were 46 undergraduates from UW–Madison who received extra credit in their Psychology course (mean age: 18.96, range: 18 to 23). They provided their demographic information in free-response text boxes. The genders of participants included 39 women, 8 men, and 1 non-binary person. Their race/ethnicity included 39 White, 3 Asian, 1 Native American, 1 Mixed, 1 African-American, 1 Dominican, and 1 Indian (please note that the number of entries was greater than the number of participants because some participants indicated more than one race/ethnicity). Participants completed this experiment on their own devices.

S.2 Finding a semantic opposite of *calm*

Our approach was based on using semantic embeddings derived from GloVe [52], a natural language processing (NLP) word embedding model. GloVe has been validated not only as a good NLP model but also as a good proxy for people’s semantic representations as measured using behavioral and neural measurements [22, 53], making it a suitable model for our purposes. Since the dimensions of GloVe need not perfectly align with the dimensions latent in people’s association ratings from Exp. 1, we needed a method to map the human principal component space to GloVe space. To do so, we estimated a transformation matrix M that mapped from the 300-D GloVe space to PC space using scipy’s `lstsq` least-squares solver. Using M , any word’s GloVe embeddings could be mapped to human PC space. We began by finding the words most dissimilar to *calm* based on their GloVe embeddings, projecting them into PC space using M and then visualizing the position of each concept. We iteratively performed this process with candidate antonyms, until we found a word that loaded on the opposite end of the PC that *calm* loaded positively on *without* loading much on either direction on the *positive – negative* PC. We arrived at *startling* as a good candidate based on this approach.

Table S.1: Hexadecimal codes for each color within the 16 color scales used in Experiments 1 and 2 (indicated by overall hue (Red, Yellow, Green, Blue), lightness (Light, Dark), and chroma (High, Low)).

Hue	Lightness	Chroma	color 1	color 2	color 3	color 4	color 5	color 6	color 7	color 8	color 9
R	L	Hi	#ffdaff	#ffb6eb	#ff91cf	#ff68b8	#ff31a3	#ff008e	#ff0078	#de0065	#bb0056
R	L	Lo	#fffcff	#f9dae4	#f4bbc8	#f39db1	#ef7f9c	#e26087	#ca4571	#ab355e	#892c4f
R	D	Hi	#eb7cb3	#da578f	#cf2975	#c80062	#be004f	#ad003d	#92002b	#73001c	#55000a
R	D	Lo	#a59fad	#9b8089	#96626f	#93455b	#8d2449	#7f0037	#680024	#4c0014	#340000
Y	L	Hi	#ffff8d	#f9ed6b	#e5da48	#d0c81b	#beb600	#aca500	#9a9500	#8a8500	#7b7600
Y	L	Lo	#fffef3	#f0ebd2	#ddd8b3	#ccc695	#bbb379	#aca25e	#9c9245	#8d832c	#7f7413
Y	D	Hi	#857c00	#6f6c00	#5c5b00	#4c4c00	#3f3c00	#352e00	#2f2000	#2b1400	#250100
Y	D	Lo	#7f7a70	#6d6953	#5d5939	#4d4920	#3e3b04	#312d00	#281f00	#211300	#1a0000
G	L	Hi	#83fff3	#15ffe2	#00ead0	#00d8c0	#00c6b1	#00b19d	#009786	#007f6f	#006856
G	L	Lo	#f7fff3	#c7eae2	#93d7d0	#54c8c0	#00b9b0	#00a59c	#008e85	#00766e	#005f55
G	D	Hi	#00b896	#009f87	#008a77	#007969	#00675b	#00524a	#003f36	#003221	#002809
G	D	Lo	#9aa496	#6d8e87	#397c77	#006e68	#005f5a	#004c49	#003835	#002821	#001a09
B	L	Hi	#c7ffff	#97e8ff	#46d3ff	#00c3ff	#00b3ff	#00a1ff	#008aff	#0074fe	#005ed8
B	L	Lo	#feffff	#dae3f5	#afceef	#7dbeee	#3baeea	#009bdf	#0085c8	#006ea9	#005987
B	D	Hi	#62a6f9	#108cee	#007ae6	#006ce4	#005ddf	#004dd4	#003abc	#00299d	#00197b
B	D	Lo	#a0a1a2	#7f8898	#567592	#156691	#00588d	#004783	#00336e	#002253	#001035

Table S.2: The mean a^* , b^* , C^* , and L^* values for each of the 16 color scales.

Hue	Lightness	Chroma	a^*	b^*	C^*	L^*
R	L	Hi	1.93434247	-0.1454748361	63.72081	63.10979
R	L	Lo	1.04967609	0.0901855282	35.76330	65.88214
R	D	Hi	1.73771836	0.3027289812	58.28588	39.65978
R	D	Lo	0.87749921	0.1607637458	31.44501	34.33223
Y	L	Hi	-0.54932776	2.1278762379	65.95778	72.95918
Y	L	Lo	-0.30369870	0.9886564320	30.15460	73.13807
Y	D	Hi	-0.19653845	1.0734107631	34.35054	25.98140
Y	D	Lo	-0.16100056	0.5724112557	17.91222	25.22744
G	L	Hi	-1.55429718	0.0032070844	42.82724	69.30993
G	L	Lo	-1.03120908	-0.0694310119	27.49006	67.42741
G	D	Hi	-1.14410813	0.1328061352	31.09054	38.28996
G	D	Lo	-0.73006326	0.0009802904	18.57621	35.20695
B	L	High	-0.23798045	-1.4679156558	50.74219	69.05558
B	L	Low	-0.30639061	-0.9043147546	29.10485	67.42697
B	D	Low	-0.04389614	-0.8139315872	26.44173	35.41818
B	D	High	0.65927419	-2.0519586083	69.03252	41.64495

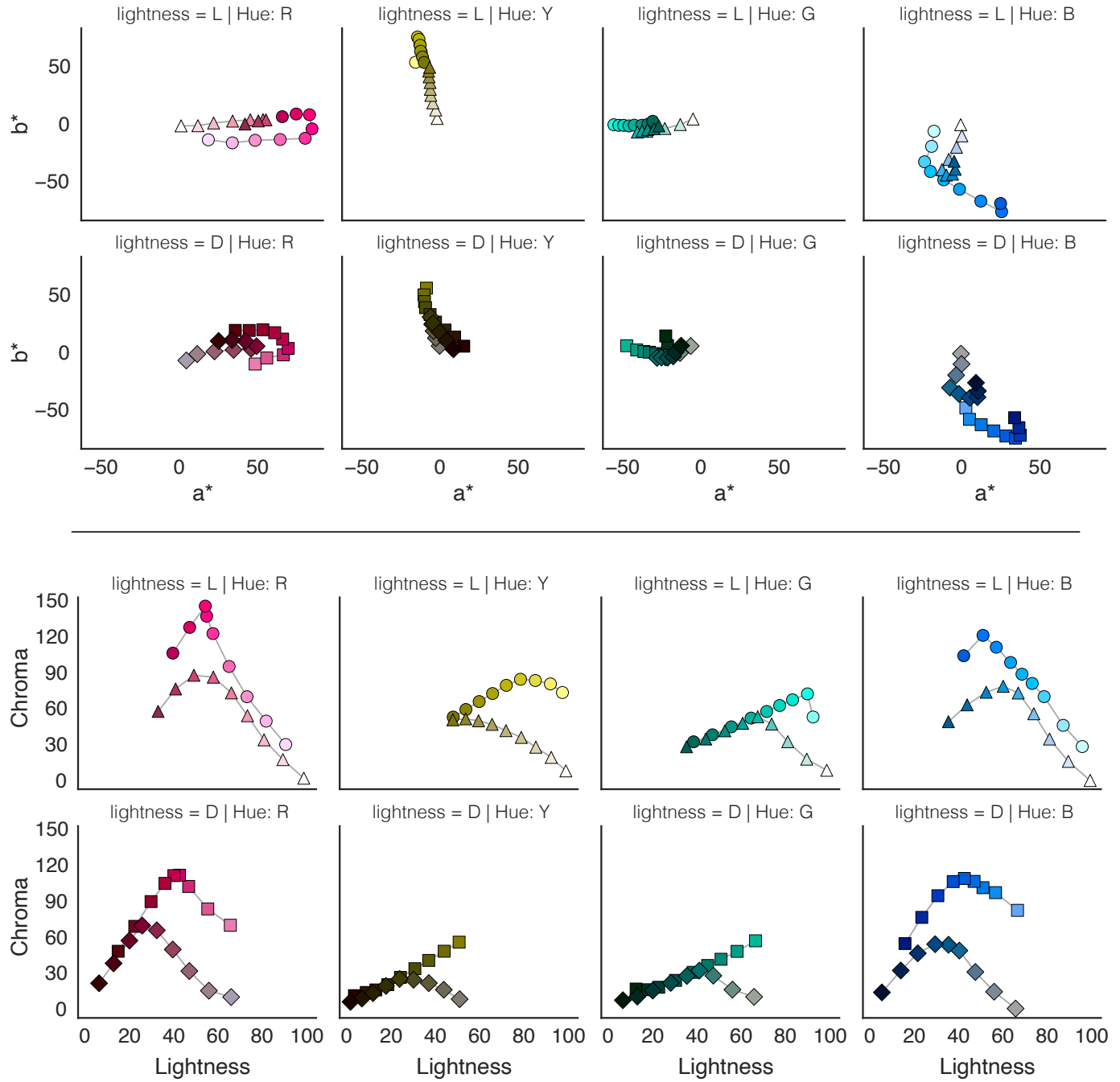


Fig. S.1: Each of the 16 color scales plotted according to their a^* (roughly red/green) vs. b^* (roughly yellow/blue) coordinates in CIELAB space (top) and according to their L^* (lightness) and C^* (chroma) coordinates in CIELCh space (bottom). Note: ' L^* ' is the same in CIELAB and CIELCh spaces. Each facet depicts both the high chroma and low chroma color scales within a given lightness level with colors belonging to the same color scale being connected by a line. The coordinates shown here appear different from the coordinates plotted within the Color Crafter tool, with discrepancies likely due to clipping and rounding issues not visualized in the tool. Marker shapes use the same mappings used in Exp. 1 (Fig. 5) and Exp. 2 (Fig. 9). Circles: light+high chroma, triangles: light+low chroma, squares: dark+high chroma, diamonds: dark+low chroma.

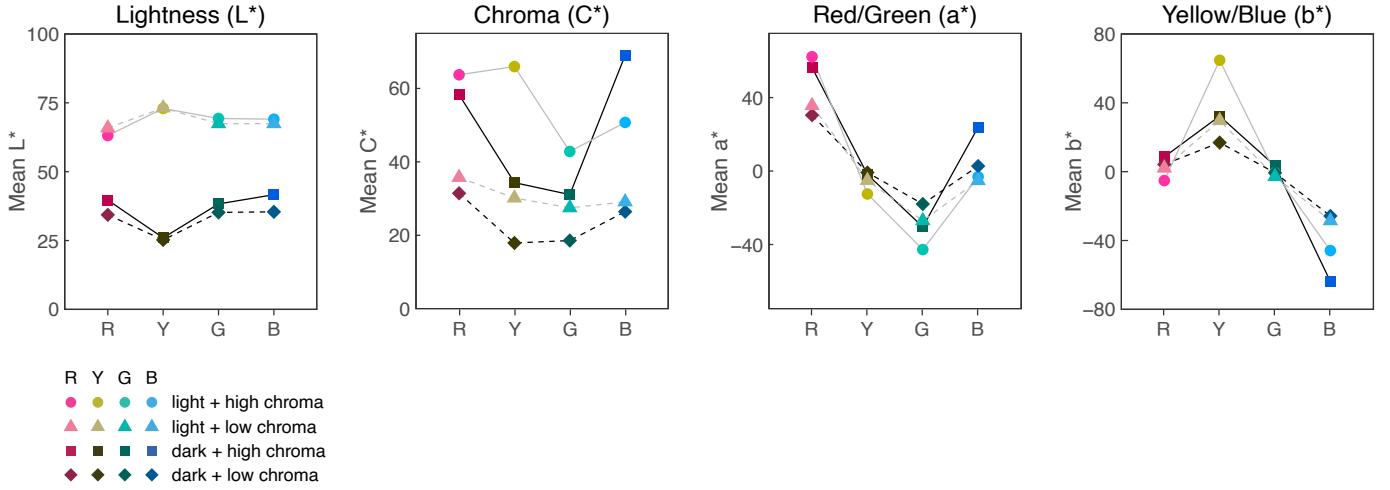


Fig. S.2: The mean L*, C*, a*, and b* coordinates for each color scale as used in the analyses of the association data, plotted in the same way as the association data shown in Figs. 5 and 9. The values correspond to the coordinates in Table S.2.

Table S.3: The mean a*, b*, C*, and L* values across colormaps constructed from each of 16 color scales applied to the light or dark shifted underlying datasets.

Hue	Lightness	Chroma	Shift	a*	b*	C*	L*
R	L	Hi	lightshift	40.0135614	-10.6921325	42.87100	78.20307
R	L	Hi	darkshift	67.5073924	2.6699313	68.19989	48.63701
R	L	Lo	lightshift	14.7269739	0.1099563	14.96044	86.71514
R	L	Lo	darkshift	42.4642460	1.4143864	42.51972	43.47099
R	D	Hi	lightshift	54.0810480	-2.0704807	55.19350	54.97053
R	D	Hi	darkshift	45.3967416	16.3722232	49.18207	24.84947
R	D	Lo	lightshift	14.7977216	-2.4443576	16.81123	55.12965
R	D	Lo	darkshift	29.7242735	8.0089109	31.25016	18.29705
Y	L	Hi	lightshift	-14.0808327	59.7983622	61.50356	88.30894
Y	L	Hi	darkshift	-10.7908360	57.9009221	58.90664	56.06223
Y	L	Lo	lightshift	-2.7583008	13.2590314	13.56206	91.11995
Y	L	Lo	darkshift	-6.0373247	41.8980402	42.34626	58.31579
Y	D	Hi	lightshift	-6.1480114	45.7094150	46.84775	39.77389
Y	D	Hi	darkshift	9.6213069	14.4209140	21.60028	10.63624
Y	D	Lo	lightshift	-0.9956447	12.3967659	12.81686	39.63490
Y	D	Lo	darkshift	5.0790834	8.8927506	12.99972	10.06670
G	L	Hi	lightshift	-41.6368678	-3.1385800	41.84410	86.28464
G	L	Hi	darkshift	-35.3379998	1.2924016	35.42387	50.46052
G	L	Lo	lightshift	-14.4266199	1.1910087	15.63487	87.17643
G	L	Lo	darkshift	-27.6476810	-1.9855748	27.82384	45.28313
G	D	Hi	lightshift	-40.7530573	4.5757776	41.16170	56.36026
G	D	Hi	darkshift	-23.8534733	10.4018087	26.83526	19.70521
G	D	Lo	lightshift	-11.6968128	2.8149495	13.28014	53.27698
G	D	Lo	darkshift	-15.3398031	2.7360430	16.33281	16.31158
B	L	Hi	lightshift	-13.5420398	-22.0317963	31.21908	85.40398
B	L	Hi	darkshift	15.9520056	-62.5756042	66.78684	51.98345
B	L	Lo	lightshift	-2.1692914	-11.8800853	12.23605	87.02787
B	L	Lo	darkshift	-5.0518189	-32.3520955	32.79897	46.64972
B	D	Hi	lightshift	10.4725279	-54.9298526	56.67350	57.75702
B	D	Hi	darkshift	31.1647432	-60.3134036	68.42741	26.22301
B	D	Lo	lightshift	0.7338501	-12.6481391	12.96321	51.99386
B	D	Lo	darkshift	7.5448208	-27.5347422	28.89766	15.63829

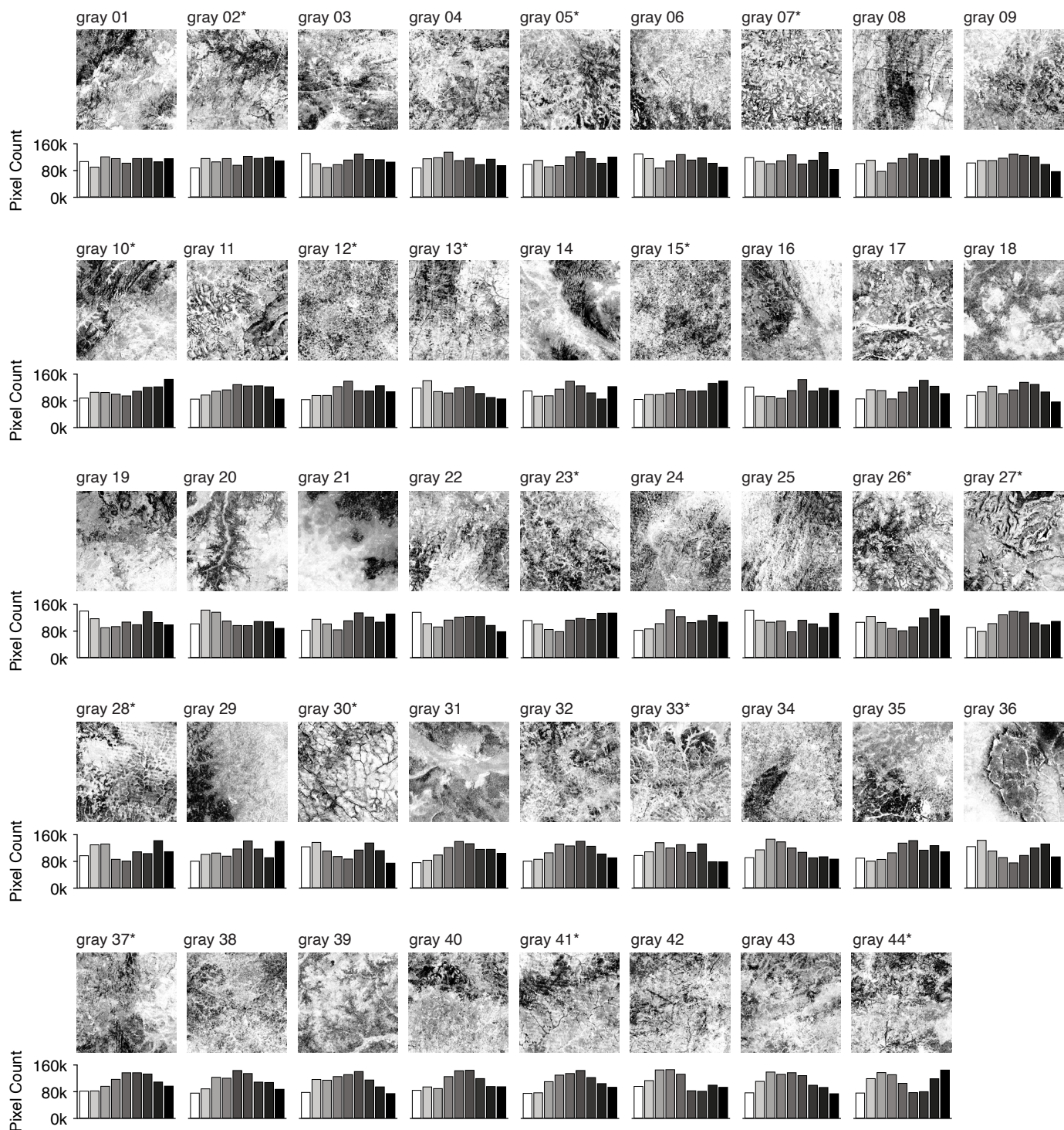


Fig. S.3: Grayscale colormaps of the 44 subtiles selected for the Exp. 1 pilot. The colormaps show biomass density values scaled to range from 0 to 100, grouped into 9 bins, and then mapped to a 9-step gray color scale (see main text for details). The histogram below each colormap shows the number of pixels in each of the 9 bins, with each bar color in the histogram corresponding to the pixel colors assigned to values in that bar's bin. The maps are labeled 01 to 44 in order from most to least uniform distribution within the set. An asterisk next to a map number indicates that map was selected for Exp. 1.

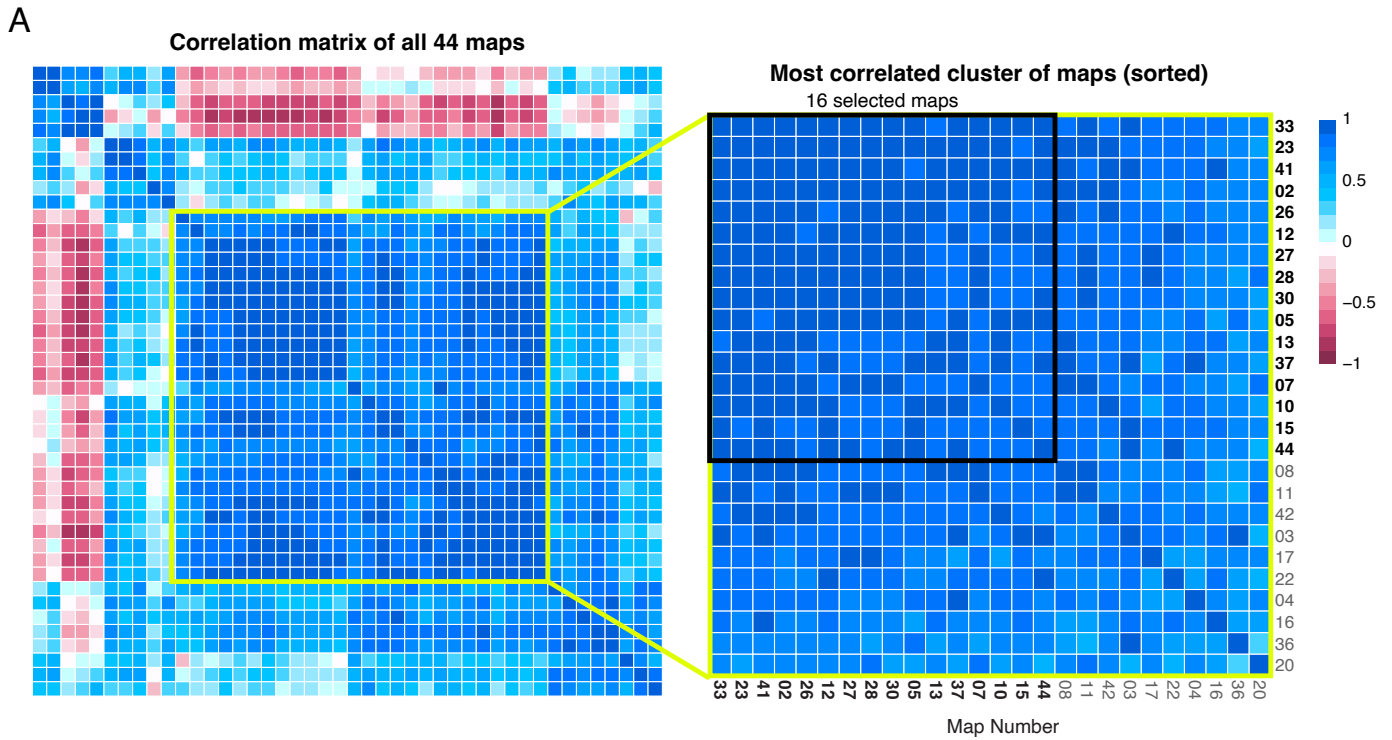


Fig. S.4: (A) Pairwise correlations between all 44 grayscale maps in terms of their mean associations with each of the 8 emotion concepts. The inset highlights the subset of 26 highly correlated maps with the black border indicating the final set of 16 selected maps. (B) Mean association ratings between each emotion concept and the 44 grayscale maps sorted from high to low. Black bars correspond to the 16 selected maps.

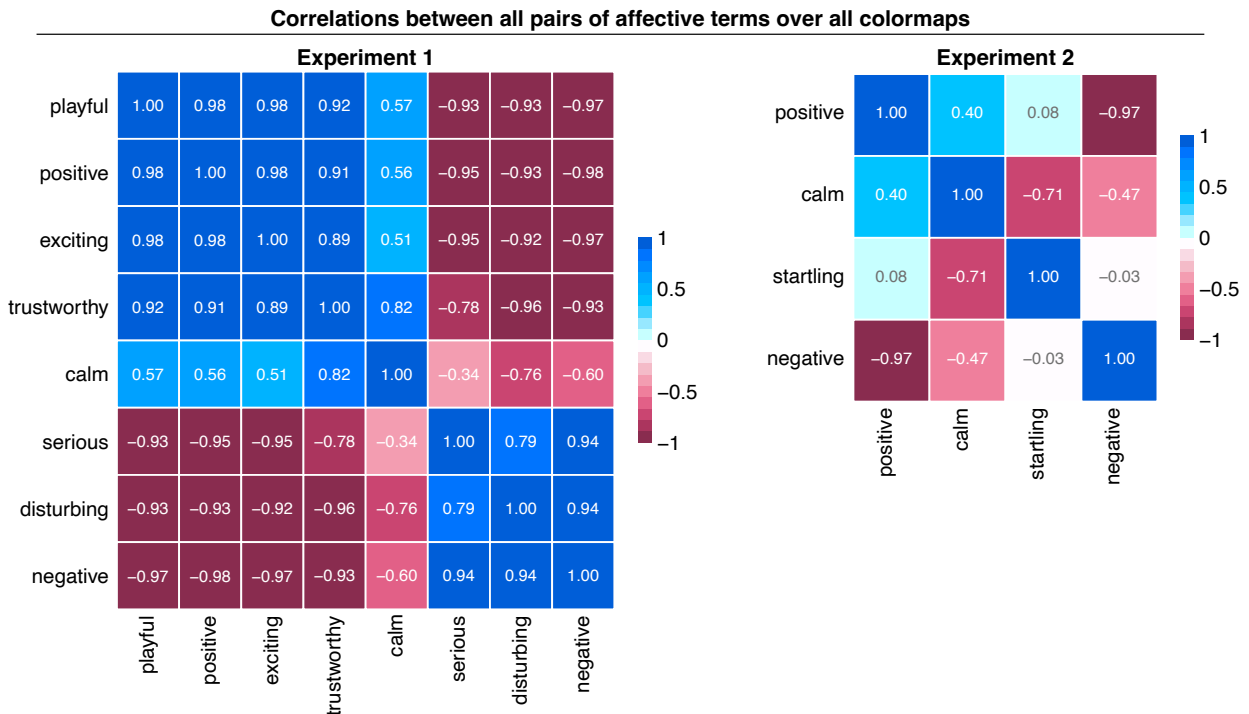


Fig. S.5: Correlation for each pairwise combination of affective terms across all colormaps in Experiment 1 (left) and Experiment 2 (right).

Table S.4: Experiment 1 mixed effects model results. Predictors were z -scored. (Stars are omitted from L*, C*, a*, and b* for table readability.)

Positive						Exciting					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.532	40.9	0.021	25.87	0.0000	(Intercept)	0.530	41.2	0.021	25.30	0.0000
L	0.116	39.0	0.021	5.47	0.0000	L	0.120	42.5	0.019	6.34	0.0000
C	0.059	104.0	0.012	5.08	0.0000	C	0.070	72.1	0.014	4.85	0.0000
a	-0.029	43.6	0.014	-2.06	0.0458	a	-0.014	64.7	0.010	-1.37	0.1745
b	-0.067	44.4	0.014	-4.83	0.0000	b	-0.080	47.6	0.013	-6.07	0.0000
L:C	-0.019	861.0	0.009	-2.07	0.0386	L:C	-0.012	857.0	0.009	-1.29	0.1958
L:a	-0.017	861.0	0.009	-2.00	0.0456	L:a	-0.019	857.0	0.009	-2.16	0.0310
L:b	-0.015	861.0	0.007	-2.15	0.0317	L:b	-0.038	857.0	0.007	-5.32	0.0000
C:a	0.034	861.0	0.008	4.19	0.0000	C:a	0.043	857.0	0.008	5.12	0.0000
C:b	0.021	861.0	0.006	3.23	0.0013	C:b	0.033	857.0	0.007	5.03	0.0000
L:C:a	0.008	861.0	0.009	0.83	0.4086	L:C:a	-0.000	857.0	0.010	-0.05	0.9622
L:C:b	-0.002	861.0	0.007	-0.35	0.7248	L:C:b	-0.009	857.0	0.007	-1.30	0.1936
Playful						Negative					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.511	48.0	0.018	27.64	0.0000	(Intercept)	0.438	39.4	0.023	19.12	0.0000
L	0.132	40.5	0.022	6.06	0.0000	L	-0.109	43.7	0.018	-5.99	0.0000
C	0.044	78.1	0.015	2.99	0.0037	C	-0.049	71.1	0.014	-3.49	0.0008
a	-0.003	49.9	0.013	-0.22	0.8237	a	0.026	40.3	0.016	1.68	0.1009
b	-0.084	45.2	0.015	-5.65	0.0000	b	0.060	42.9	0.015	4.00	0.0002
L:C	-0.008	857.0	0.010	-0.78	0.4342	L:C	0.027	853.0	0.009	2.88	0.0041
L:a	-0.022	857.0	0.010	-2.25	0.0247	L:a	-0.015	853.0	0.009	-1.64	0.1017
L:b	-0.017	857.0	0.008	-2.21	0.0272	L:b	0.016	853.0	0.007	2.29	0.0223
C:a	0.039	857.0	0.009	4.35	0.0000	C:a	-0.022	853.0	0.008	-2.55	0.0110
C:b	0.012	857.0	0.007	1.74	0.0817	C:b	-0.019	853.0	0.007	-2.78	0.0056
L:C:a	0.006	857.0	0.010	0.60	0.5487	L:C:a	0.010	853.0	0.010	0.98	0.3284
L:C:b	-0.005	857.0	0.008	-0.70	0.4856	L:C:b	0.008	853.0	0.007	1.04	0.2973
Disturbing						Serious					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.438	40.0	0.023	18.82	0.0000	(Intercept)	0.527	38.6	0.024	22.11	0.0000
L	-0.062	48.3	0.017	-3.64	0.0007	L	-0.144	40.4	0.021	-6.95	0.0000
C	-0.039	79.1	0.014	-2.74	0.0076	C	-0.036	103.0	0.012	-2.88	0.0048
a	0.044	46.0	0.014	3.27	0.0020	a	-0.013	55.6	0.011	-1.17	0.2473
b	0.098	40.5	0.017	5.71	0.0000	b	0.029	44.0	0.015	1.97	0.0553
L:C	0.024	848.0	0.010	2.39	0.0172	L:C	-0.001	860.0	0.009	-0.08	0.9358
L:a	-0.004	848.0	0.010	-0.40	0.6904	L:a	0.011	860.0	0.009	1.20	0.2319
L:b	0.011	848.0	0.007	1.46	0.1437	L:b	0.030	860.0	0.007	4.19	0.0000
C:a	-0.036	848.0	0.009	-4.11	0.0000	C:a	-0.026	860.0	0.008	-3.13	0.0018
C:b	-0.035	848.0	0.007	-5.05	0.0000	C:b	-0.017	860.0	0.007	-2.55	0.0111
L:C:a	0.003	848.0	0.010	0.33	0.7389	L:C:a	0.002	860.0	0.010	0.16	0.8768
L:C:b	0.004	848.0	0.008	0.51	0.6104	L:C:b	0.002	860.0	0.007	0.22	0.8259
Calm						Trustworthy					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.513	49.4	0.019	27.55	0.0000	(Intercept)	0.509	40.9	0.021	24.51	0.0000
L	0.035	38.9	0.025	1.42	0.1627	L	0.075	42.3	0.019	4.02	0.0002
C	-0.015	69.8	0.016	-0.95	0.3435	C	0.012	79.0	0.013	0.90	0.3714
a	-0.036	49.1	0.013	-2.73	0.0088	a	-0.014	48.7	0.012	-1.20	0.2366
b	-0.104	45.2	0.016	-6.70	0.0000	b	-0.091	41.7	0.015	-5.97	0.0000
L:C	-0.013	846.0	0.010	-1.22	0.2224	L:C	-0.020	854.0	0.009	-2.15	0.0321
L:a	0.005	846.0	0.010	0.53	0.5944	L:a	-0.002	854.0	0.009	-0.27	0.7842
L:b	0.011	846.0	0.008	1.45	0.1484	L:b	-0.005	854.0	0.007	-0.73	0.4635
C:a	-0.002	846.0	0.009	-0.19	0.8479	C:a	0.014	854.0	0.008	1.74	0.0829
C:b	0.005	846.0	0.007	0.70	0.4828	C:b	0.014	854.0	0.007	2.14	0.0327
L:C:a	-0.024	846.0	0.011	-2.23	0.0263	L:C:a	-0.008	854.0	0.009	-0.87	0.3839
L:C:b	0.014	846.0	0.008	1.70	0.0895	L:C:b	0.004	854.0	0.007	0.61	0.5401

Table S.5: Results from multiple linear regression models predicting emotion space principal component scores from color space coordinates (see main text for details). Predictors were z-scored.

Experiment	PC	Predictor	β	SE	t	p
Exp. 1	PC1	Intercept	0.0	0.182	0.00	1.000
		L*	1.902	0.211	8.99	<0.001
		C*	0.573	0.235	2.44	0.033
		a*	-0.143	0.225	-0.635	0.538
		b*	-1.407	0.190	-7.398	<0.001
	PC2	Intercept	0.0	0.109	0.00	1.000
		L*	0.3471	0.126	3.092	0.028
		C*	0.087	0.140	2.203	0.545
		a*	0.506	0.134	3.77	0.003
		b*	0.649	0.113	4.50	<0.001
Exp.2	PC1	Intercept	0.0	0.144	0.00	1.000
		L*	0.697	0.149	4.659	<0.001
		C*	0.6026	0.157	2.988	<0.001
		a*	-0.386	0.163	-2.373	0.025
		b*	-0.966	0.149	-6.483	<0.001
	PC2	Intercept	0.0	0.1269	0.00	1.000
		L*	0.709	0.1269	5.589	<0.001
		C*	0.208	0.1269	1.641	0.053
		a*	0.698	0.1794	3.888	<0.001
		b*	0.673	0.1794	3.751	<0.001

Correlations between each PC and affective term over all colormaps

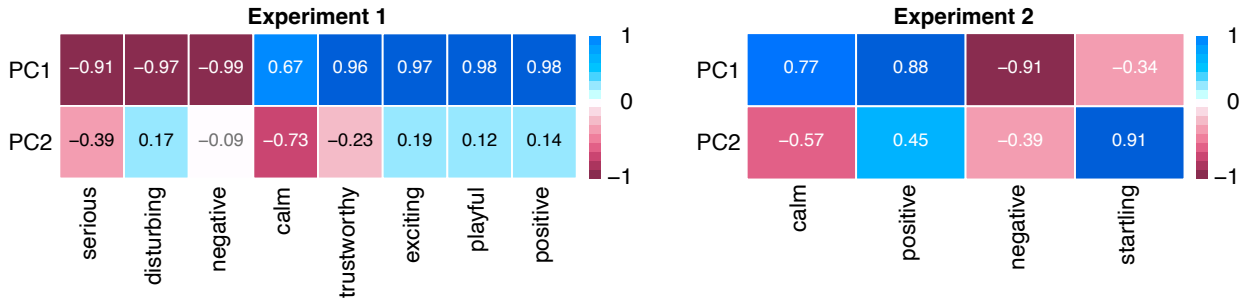


Fig. S.6: Correlations between mean affective concept ratings and PC loadings across all colormaps in Experiment 1 (left) and Experiment 2 (right).

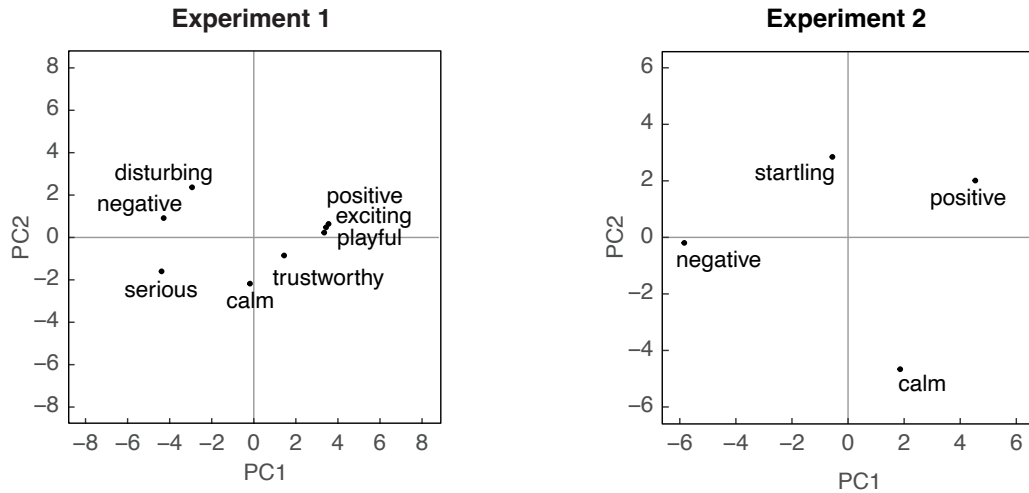


Fig. S.7: Organization of emotion concepts along the first two principal components derived from the mean **color scale** $\times$ **emotion concept** matrices in Experiment 1 (left) and Experiment 2 (right).

Table S.6: Experiment 2 mixed effects model results. Predictors were z -scored. (Stars are omitted from L*, C*, a*, and b* for table readability.)

Positive						Negative					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.516	51.9	0.013	39.38	0.0000	(Intercept)	0.475	42.9	0.017	28.48	0.0000
shift	0.046	1990.0	0.007	6.32	0.0000	shift	-0.050	1990.0	0.007	-6.95	0.0000
L	0.102	46.3	0.014	7.11	0.0000	L	-0.102	47.0	0.014	-7.53	0.0000
C	0.059	11.4	0.009	6.34	0.0000	C	-0.054	86.9	0.010	-5.33	0.0000
a	-0.007	47.6	0.011	-0.63	0.5256	a	0.036	41.3	0.013	2.75	0.0059
b	-0.064	40.8	0.014	-4.53	0.0000	b	0.080	43.7	0.012	6.51	0.0000
L:C	-0.022	1990.0	0.007	-3.01	0.0026	L:C	0.022	1990.0	0.007	3.09	0.0020
L:a	-0.004	1990.0	0.007	-0.60	0.5469	L:a	-0.009	1990.0	0.007	-1.27	0.2038
L:b	0.003	1990.0	0.006	0.52	0.6041	L:b	-0.006	1990.0	0.005	-1.06	0.2909
C:a	0.034	1990.0	0.006	5.29	0.0000	C:a	-0.044	1990.0	0.006	-6.85	0.0000
C:b	0.032	1990.0	0.005	6.18	0.0000	C:b	-0.039	1990.0	0.005	-7.63	0.0000
L:C:a	0.010	1990.0	0.007	1.34	0.1796	L:C:a	-0.006	1990.0	0.007	-0.79	0.4312
L:C:b	-0.018	1990.0	0.006	-3.32	0.0009	L:C:b	0.011	1990.0	0.005	2.02	0.0437
L:shift	-0.037	1990.0	0.007	-5.14	0.0000	L:shift	0.024	1990.0	0.007	3.40	0.0007
C:shift	0.014	1990.0	0.008	1.77	0.0771	C:shift	-0.002	1990.0	0.008	-0.29	0.7753
a:shift	-0.006	1990.0	0.006	-0.88	0.3812	a:shift	-0.005	1990.0	0.006	-0.83	0.4049
b:shift	0.014	1990.0	0.006	2.29	0.0218	b:shift	-0.019	1990.0	0.006	-3.04	0.0024
L:C:shift	-0.009	1990.0	0.007	-1.19	0.2328	L:C:shift	0.010	1990.0	0.007	1.34	0.1804
L:a:shift	-0.000	1990.0	0.007	-0.05	0.9567	L:a:shift	0.002	1990.0	0.007	0.34	0.7309
L:b:shift	0.006	1990.0	0.006	1.13	0.2584	L:b:shift	-0.008	1990.0	0.005	-1.56	0.1197
C:a:shift	-0.001	1990.0	0.006	-0.08	0.9345	C:a:shift	-0.010	1990.0	0.006	-1.61	0.1069
C:b:shift	0.003	1990.0	0.005	0.60	0.5517	C:b:shift	-0.005	1990.0	0.005	-0.98	0.3261
L:C:a:shift	-0.016	1990.0	0.007	-2.19	0.0284	L:C:a:shift	0.019	1990.0	0.007	2.55	0.0109
L:C:b:shift	-0.012	1990.0	0.006	-2.19	0.0287	L:C:b:shift	0.015	1990.0	0.005	2.70	0.0070

Calm						Startling					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.513	43.2	0.018	28.60	0.0000	(Intercept)	0.467	41.1	0.019	24.67	0.0000
shift	-0.035	1990.0	0.008	-4.46	0.0000	shift	0.010	1990.0	0.007	1.37	0.1722
L	-0.020	40.2	0.020	-1.01	0.3114	L	0.033	37.0	0.024	1.37	0.1709
C	0.030	87.5	0.011	2.74	0.0061	C	-0.000	59.6	0.013	-0.01	0.9881
a	-0.089	44.1	0.012	-7.10	0.0000	a	0.027	44.8	0.011	2.35	0.0188
b	-0.154	39.5	0.016	-9.33	0.0000	b	0.056	43.3	0.013	4.33	0.0000
L:C	-0.037	1990.0	0.008	-4.67	0.0000	L:C	0.008	1990.0	0.007	1.03	0.3016
L:a	0.011	1990.0	0.008	1.49	0.1367	L:a	-0.007	1990.0	0.007	-0.96	0.3365
L:b	0.007	1990.0	0.006	1.19	0.2337	L:b	-0.012	1990.0	0.006	-2.16	0.0310
C:a	0.024	1990.0	0.007	3.40	0.0007	C:a	-0.001	1990.0	0.007	-0.16	0.8707
C:b	0.048	1990.0	0.006	8.63	0.0000	C:b	-0.017	1990.0	0.005	-3.13	0.0018
L:C:a	0.005	1990.0	0.008	0.67	0.5057	L:C:a	0.006	1990.0	0.008	0.83	0.4039
L:C:b	-0.000	1990.0	0.006	-0.06	0.9514	L:C:b	0.003	1990.0	0.006	0.51	0.6088
L:shift	-0.008	1990.0	0.008	-1.00	0.3171	L:shift	0.033	1990.0	0.007	4.47	0.0000
C:shift	0.012	1990.0	0.009	1.40	0.1612	C:shift	-0.025	1990.0	0.008	-3.03	0.0024
a:shift	0.005	1990.0	0.007	0.75	0.4524	a:shift	-0.003	1990.0	0.007	-0.41	0.6788
b:shift	0.025	1990.0	0.007	3.83	0.0001	b:shift	-0.009	1990.0	0.006	-1.42	0.1546
L:C:shift	0.021	1990.0	0.008	2.71	0.0068	L:C:shift	-0.016	1990.0	0.007	-2.14	0.0321
L:a:shift	0.007	1990.0	0.008	0.90	0.3676	L:a:shift	-0.001	1990.0	0.007	-0.11	0.9089
L:b:shift	0.007	1990.0	0.006	1.17	0.2421	L:b:shift	-0.009	1990.0	0.006	-1.57	0.1154
C:a:shift	0.005	1990.0	0.007	0.70	0.4857	C:a:shift	0.006	1990.0	0.007	0.95	0.3406
C:b:shift	-0.016	1990.0	0.006	-2.93	0.0034	C:b:shift	0.004	1990.0	0.005	0.84	0.4031
L:C:a:shift	-0.008	1990.0	0.008	-0.93	0.3516	L:C:a:shift	-0.002	1990.0	0.008	-0.30	0.7627
L:C:b:shift	-0.006	1990.0	0.006	-0.96	0.3354	L:C:b:shift	0.017	1990.0	0.006	3.05	0.0023

Table S.7: Experiment 2 mixed effects model results using colorimetric predictors derived from color scales. Predictors were z -scored. (Stars are omitted from L*, C*, a*, and b* for table readability.)

Positive						Negative					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.516	55.9	0.013	39.54	0.0000	(Intercept)	0.475	45.6	0.017	28.68	0.0000
L	0.102	49.4	0.014	7.15	0.0000	L	-0.102	50.4	0.014	-7.53	0.0000
C	0.059	129.0	0.009	6.27	0.0000	C	-0.054	99.0	0.010	-5.28	0.0000
a	-0.007	51.3	0.011	-0.64	0.5227	a	0.036	44.1	0.013	2.79	0.0053
b	-0.064	42.9	0.014	-4.57	0.0000	b	0.080	46.5	0.012	6.56	0.0000
L:C	-0.022	2000.0	0.008	-2.90	0.0037	L:C	0.022	2000.0	0.008	2.95	0.0032
L:a	-0.004	2000.0	0.007	-0.58	0.5614	L:a	-0.009	2000.0	0.007	-1.21	0.2247
L:b	0.003	2000.0	0.006	0.50	0.6171	L:b	-0.006	2000.0	0.006	-1.01	0.3130
C:a	0.034	2000.0	0.007	5.10	0.0000	C:a	-0.044	2000.0	0.007	-6.55	0.0000
C:b	0.032	2000.0	0.005	5.96	0.0000	C:b	-0.039	2000.0	0.005	-7.29	0.0000
L:C:a	0.010	2000.0	0.008	1.29	0.1957	L:C:a	-0.006	2000.0	0.008	-0.75	0.4520
L:C:b	-0.018	2000.0	0.006	-3.20	0.0014	L:C:b	0.011	2000.0	0.006	1.93	0.0540

Calm						Startling					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.513	45.6	0.018	28.91	0.0000	(Intercept)	0.467	42.9	0.019	24.97	0.0000
L	-0.020	41.9	0.019	-1.02	0.3060	L	0.033	48.4	0.024	1.39	0.1655
C	0.030	94.3	0.011	2.74	0.0062	C	-0.000	63.2	0.013	-0.01	0.9881
a	-0.089	46.3	0.012	-7.16	0.0000	a	0.027	47.0	0.011	2.37	0.0177
b	-0.154	41.2	0.016	-9.45	0.0000	b	0.056	45.3	0.013	4.38	0.0000
L:C	-0.037	2010.0	0.008	-4.59	0.0000	L:C	0.008	2010.0	0.008	1.02	0.3087
L:a	0.011	2010.0	0.008	1.46	0.1435	L:a	-0.007	2010.0	0.007	-0.95	0.3435
L:b	0.007	2010.0	0.006	1.17	0.2418	L:b	-0.012	2010.0	0.006	-2.13	0.0336
C:a	0.024	2010.0	0.007	3.34	0.0008	C:a	-0.001	2010.0	0.007	-0.16	0.8726
C:b	0.048	2010.0	0.006	8.48	0.0000	C:b	-0.017	2010.0	0.005	-3.08	0.0021
L:C:a	0.005	2010.0	0.008	0.65	0.5130	L:C:a	0.006	2010.0	0.008	0.82	0.4108
L:C:b	-0.000	2010.0	0.006	-0.06	0.9522	L:C:b	0.003	2010.0	0.006	0.50	0.6140

Table S.8: Experiment 2 mixed effects model results using colorimetric predictors derived from colormap images. Predictors were z -scored. (Stars are omitted from L*, C*, a*, and b* for table readability.)

Positive						Negative					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.510	38.2	0.012	43.10	0.0000	(Intercept)	0.477	36.2	0.016	30.51	0.0000
L	0.113	38.3	0.015	7.46	0.0000	L	-0.114	37.9	0.015	-7.39	0.0000
C	0.080	44.4	0.009	8.97	0.0000	C	-0.072	43.0	0.009	-7.87	0.0000
a	-0.017	42.8	0.010	-1.65	0.0980	a	0.044	39.5	0.013	3.56	0.0004
b	-0.101	52.0	0.015	-6.85	0.0000	b	0.114	60.6	0.013	9.05	0.0000
L:C	-0.015	2000.0	0.006	-2.60	0.0094	L:C	0.016	2000.0	0.006	2.83	0.0047
L:a	0.013	2000.0	0.006	2.15	0.0317	L:a	-0.026	2000.0	0.006	-4.46	0.0000
L:b	0.008	2000.0	0.008	1.07	0.2854	L:b	-0.014	2000.0	0.007	-1.94	0.0528
C:a	0.044	2000.0	0.006	7.21	0.0000	C:a	-0.050	2000.0	0.006	-8.37	0.0000
C:b	0.054	2000.0	0.006	8.56	0.0000	C:b	-0.056	2000.0	0.006	-9.18	0.0000
L:C:a	-0.010	2000.0	0.010	-1.03	0.3024	L:C:a	0.002	2000.0	0.010	0.16	0.8734
L:C:b	-0.025	2000.0	0.006	-4.07	0.0000	L:C:b	0.014	2000.0	0.006	2.36	0.0184

Calm						Startling					
Predictor	β	df	SE	t	p	Predictor	β	df	SE	t	p
(Intercept)	0.502	36.2	0.017	29.94	0.0000	(Intercept)	0.463	35.8	0.018	25.93	0.0000
L	-0.042	36.1	0.022	-1.91	0.0563	L	0.042	35.6	0.024	1.71	0.0876
C	0.045	41.8	0.010	4.34	0.0000	C	-0.012	38.5	0.013	-0.95	0.3438
a	-0.103	40.5	0.012	-8.45	0.0000	a	0.034	40.9	0.012	2.90	0.0037
b	-0.186	47.2	0.017	-10.93	0.0000	b	0.070	56.1	0.014	4.99	0.0000
L:C	-0.032	2010.0	0.006	-5.27	0.0000	L:C	0.008	2010.0	0.006	1.38	0.1686
L:a	0.018	2010.0	0.006	2.85	0.0043	L:a	-0.019	2010.0	0.006	-3.17	0.0015
L:b	0.003	2010.0	0.008	0.39	0.6972	L:b	-0.015	2010.0	0.008	-1.93	0.0535
C:a	0.031	2010.0	0.006	4.95	0.0000	C:a	0.001	2010.0	0.006	0.10	0.9184
C:b	0.071	2010.0	0.006	11.02	0.0000	C:b	-0.030	2010.0	0.006	-4.63	0.0000
L:C:a	-0.007	2010.0	0.010	-0.73	0.4642	L:C:a	-0.001	2010.0	0.010	-0.07	0.9476
L:C:b	-0.000	2010.0	0.006	-0.05	0.9568	L:C:b	0.008	2010.0	0.006	1.28	0.2001