Quantifying Emotional Responses to Immutable Data Characteristics and Designer Choices in Data Visualizations

Carter Blair, Xiyao Wang, and Charles Perin

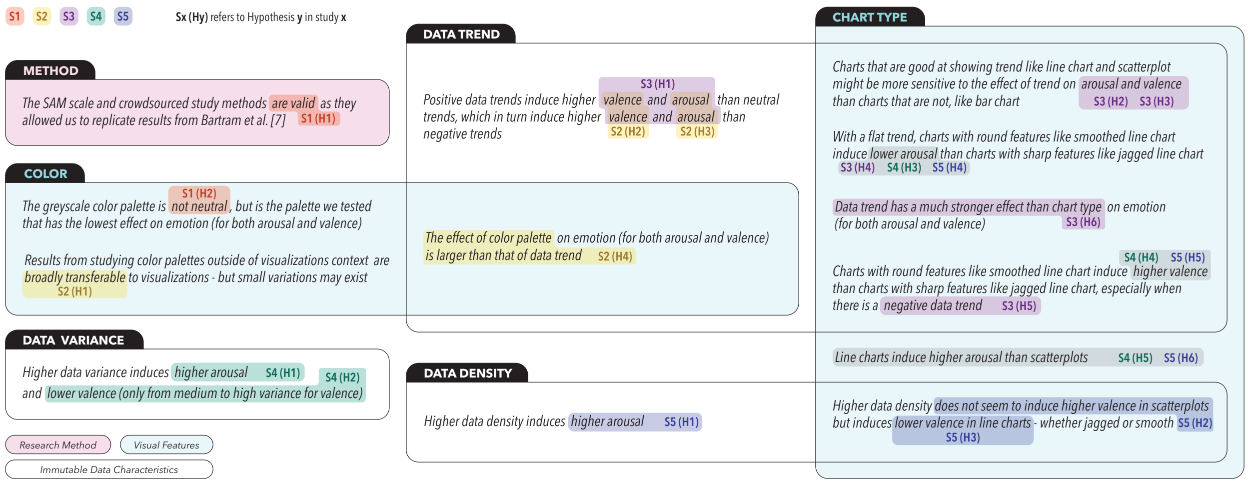

Fig. 1: Summary of results from the five studies S1–S5.

Abstract—Emotion is an important factor to consider when designing visualizations as it can impact the amount of trust viewers place in a visualization, how well they can retrieve information and understand the underlying data, and how much they engage with or connect to a visualization. We conducted five crowdsourced experiments to quantify the effects of color, chart type, data trend, data variability and data density on emotion (measured through self-reported arousal and valence). Results from our experiments show that there are multiple design elements which influence the emotion induced by a visualization and, more surprisingly, that certain data characteristics influence the emotion of viewers even when the data has no meaning. In light of these findings, we offer guidelines on how to use color, scale, and chart type to counterbalance and emphasize the emotional impact of immutable data characteristics.

Index Terms—Affect, Data Visualization, Emotion, Quantitative Study

1 Introduction

Emotion is an important factor to consider when designing data visualizations. Viewer's emotion can impact the extent to which they engage with or connect to a visualization [15, 30, 34] as well as their trust in [39], and understanding of [22], the data. Emotional response matters in many decision-making and communication scenarios: a healthcare provider might want a patient to stay calm when presented with medical information, a media might aim to arouse viewers, or a business might aim to get investors excited about their financial results.

The field of psychology has established that visual stimuli can induce emotion [27, 58]. It has been found that this is also true for both static [7] and dynamic [30, 33, 41] data visualizations. However, empirical evaluations are missing to quantify the effects of visualization features other than color [7] on emotion. This is an important gap to address so that visualization designers can make informed design choices to counterbalance and enhance the emotional impact arising from immutable data characteristics in visualizations.

• *Carter Blair is with the University of Waterloo. E-mail: cblair@uwaterloo.ca*
• *Xiyao Wang is with the University of Victoria and TU Delft. E-mail: xiyao.wang23@gmail.com*
• *Charles Perin is with the University of Victoria. E-mail: cperin@uvic.ca*

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.xxxx/TVCG.201x.xxxxxxx

We show that **chart type**, **data trend**, **data variance**, and **data density** all influence the emotion induced by a visualization—in addition to **color** [7]. We quantify through five studies the effects of color (S1 and S2), chart type (S3, S4, and S5), data trend (S2 and S3), data variance (S4), and data density (S5) on emotion (measured through arousal and valence ratings using the Self-Assessment Manikin scale [35]).

Remarkably, this work is the first to show that ***even if the data has no semantic meaning, certain data characteristics influence the emotion of viewers***. Figure 1 summarizes the results from the five studies. It shows that two fundamental designer choices (color and chart type) and three fundamental immutable data characteristics (data trend, variance, and density) influence viewers' emotions regardless of a visualization's semantic meaning. From these results, we provide guidelines on how to manipulate color, scale, and chart type to counterbalance or emphasize the emotional impact of immutable data characteristics. This is a starting point for fully understanding the effect of visualization features on viewers' emotions.

2 Related Work

Work in psychology and visualization sheds light on potential effects of designer choices and immutable data characteristics on emotions.

2.1 The role of emotion in visualization

Emotional response should be considered to design and assess the value of visualizations [18, 32, 66] for several reasons: emotional priming with positive emotion-inducing stimuli increases the accuracy of information extraction from visualizations [22]; negative emotions have controversial effects and can sometimes increase long-term recall [34]; emotion can affect engagement with data [30]; and there is a relationship between color, perceived beauty, and the degree to which a visualization is trusted by viewers [39]. In data storytelling also, creating narratives that induce emotion is an important consideration [56] to help viewers engage with and relate to the information [15].

Together, these factors emphasize the significance of investigating the elements of visualizations that impact emotion and their extent.

2.2 Measuring emotion

Measures of emotion can be behavioral, physiological, or experiential [14, 47]. Behavioral measures quantify overt acts and secondary behaviors such as voice pitch (e. g., [3]) or facial behavior (e. g., [59]). Physiological measures (e. g., [26]) focus on physiological reactions such as changes in brain states or skin conductance. Experiential measures (e. g., [42]) quantify self-reported subjective experiences. *Behavioral* and *physiological* measures require moderation because they rely on dedicated equipment such as audio and video recordings and physiological sensors. We used *experiential* measures because i) self-assessment can be administered as a survey question, and ii) they are reliable to measure arousal and valence in crowdsourced studies [8].

Emotion is commonly characterized through three factors: arousal, valence, and dominance [13]. Given a stimulus, arousal corresponds to the intensity of emotion it induces, valence to its pleasantness and dominance to its perceived power [68]. Because arousal and valence are more significant and discriminatory [13, 50] and more reliable [13, 47] measures than dominance, we used the SAM rating system [35] to measure arousal and valence on nine-point scales (see Figure 2).

2.3 The effect of color on emotions

Color has been the most studied feature when it comes to emotion. Its role has been largely explored in psychology. It has been shown, for example, that the printed paper color has an influence on emotions when reading [69]; that saturated and bright colors are associated with higher arousal and valence and that hue has an effect on both measures [70]; that principle hues (such as red, yellow, and green) elicit more positive responses than intermediate hues (such as yellow-red and blue-green), which in turn elicit more positive responses than achromatic colors such as grey and black [28]; and that bright colors are associated with high valence and saturated colors with high arousal [64].

The role of color on emotion has also received attention in visualization; unsurprisingly, given the importance of color in visualization design (e. g., [25, 40, 51, 57, 60, 62, 67]). Bartram et al. [7], building on results from psychology, studied the relationship between color in bar charts and emotion, from which they designed a series of palettes that convey different emotions. Anderson et al. [1] further validated the emotional impact of varying color palettes when interpreting categorical maps. A recent study on arousal-related design features also highlighted the affective influence of color on visualizations [33].

2.4 What other features could affect emotions?

Other features than color have an effect on emotion. Motion, for example, affects emotion [6, 41]; and for a static visualization, the trajectory of the data can provide a sense of motion. Shape can also affect emotion, given that humans tend to prefer rounded objects over sharp objects [4], with sharp objects having high arousal and low valence [2]. Leder et al. [37] distinguished positive (like cake) and negative (like snake) stimuli and found that people's preference towards curved contours over sharp ones held for positive or neutral stimuli, but no difference was identified for negative stimuli. Various other studies in psychology have shown that people's preferences vary with different shape properties, such as symmetry [63], size [61], and complexity [55].

The chart type influences the perception of trend and determines the shape of the visualization, affecting users' preference [65] and memorability [11]. Interestingly, motion and shape can be manipulated by a visualization designer (e. g., using motion to encode data and using circles vs. squares to represent data points in a scatterplot), but they are also influenced by immutable data characteristics such as trend, variance and density, that create the direction of motion.

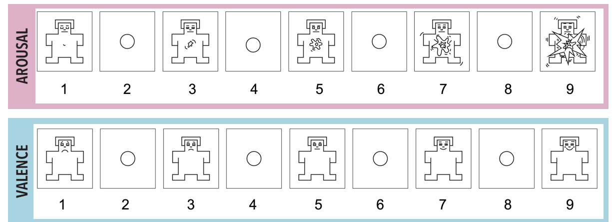

Fig. 2: SAM scales used in our studies (figures from [10]).

Based on the above, and acknowledging that many features and data characteristics could be studied, we identified a first set of visualization aspects to study. In addition to color, we include chart type, data trend, data variance, and data density. We chose these because i) they are potentially emotion-inducing features, ii) together they allow us to investigate both features that are due to immutable data characteristics that designers cannot manipulate and features that visualization designers can manipulate, and iii) we lack empirical data about the relationship between these characteristics and features, and emotion.

3 Study Rationale and Methodology

The aim of this research is to examine the relationship between emotion and three immutable data characteristics (data trend, data variance and data density), and two visual features of visualizations (color and chart type). To achieve this, we conducted five crowdsourced studies (S1–S5, approved by our institutional Ethics board) that asked participants to rate their self-perceived arousal and valence in response to visualizations. We study color (S1); color and data trends (S2); data trends and chart types (S3); data variance and chart types (S4); and data density and chart types (S5). We chose those factors due to their ubiquity in visualization, and we studied at most two factors per study to keep the participant time requirements reasonable. Studies 2, 3, 4, and 5 each had two independent variables—one a design choice and the other a data characteristic. Next, we describe the methodology common to all studies (the specificity of each is provided in their respective sections).

Measurements. We used the SAM rating system [35] to measure arousal and valence on nine-point scales (see Figure 2 and Section 2.2). In the remainder of the paper, A refers to the arousal ratings for A and A to the valence ratings for A. $A > B$ means that the arousal for A is larger than for B, and $A > B$ that the valence for A is larger than for B.

Design. Studies are within-subject designs in which all participants completed the same tasks for all conditions. The studies are divided into measurement blocks, each containing all conditions for one measurement. The number of blocks depends on the number of repetitions and is counter-balanced. The order of conditions within each block is randomized. The studies were created with both *SurveyMonkey* and *Qualtrics* and were deployed through the *Prolific* crowdsourcing platform. Figure 3 shows the structure of the studies.

Procedure. Each participant completed the study online without supervision. After selecting the study on Prolific, they were redirected to the study survey on SurveyMonkey/Qualtrics. After participants had read the consent form and given consent to participate, they were given illustrated explanations about the rating scales (arousal and valence) and instructions on how to rate their emotions with these scales. Then, they answered each question one by one – they only saw the next question after the current one was confirmed, and they were not allowed to go back and change their answer.

Participants. We recruited participants online through Prolific. They had to be over 18 years old, have normal or corrected to normal vision without color-deficiency, and have a Prolific approval rating over 99% (to ensure quality responses). They were compensated at least minimum hourly wage based on the estimated duration of each study. We excluded participants who spent a total time far shorter than the estimated study time or who input the same answers for most conditions. We provide study participant demographics in supplemental materials.

Analysis. Each study (aside from S1) had two independent variables v_1 and v_2 and two measures. To assess the effects of v_1 and v_2, we used a 3-step procedure: for each participant p and measure m, 1) Averaged p's responses to all repetitions of a given $\{p, \text{condition}, m\}$, using the mean;

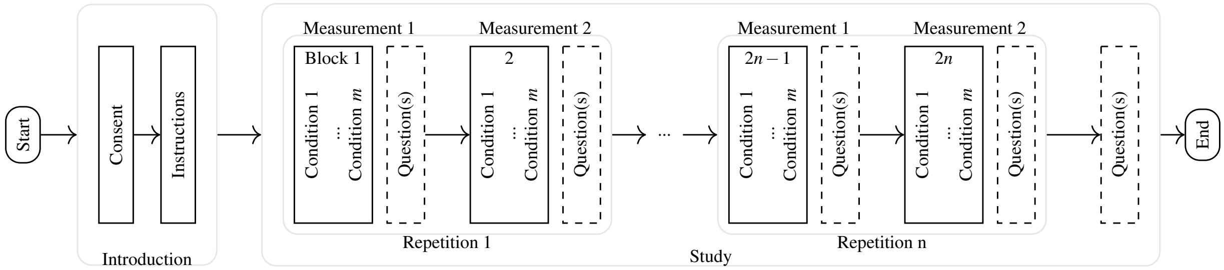

Fig. 3: Workflow for the five studies. Dashed blocks are optional question(s) asked depending on each study design. Measurements 1 and 2 are arousal or valence, depending on the counter-balanced order.

2) Calculated aggregate values for each level of v_1 and v_2 by averaging across all levels of the other variable. 3) Generated confidence intervals (CIs) by either: a) Bootstrapping aggregate values from all participants, or b) Bootstrapping differences between levels from all participants. There are two exceptions to this process: 1) in S1, there was only one variable, so we skipped step 2; 2) for hypotheses that concern the variance induced by an independent variable (S2(H4) and S3(H6)), we used standard deviation, as explained in Section 5.3. We interpret the results using effect sizes and CIs rather than $p-$values, following recommendations for statistical practices in HCI and visualization (e. g., [17, 19]). We report the mean and 95% CI values for each result.

4 S1 – Color Palettes

Study 1 investigates people's emotional responses to five color palettes. The first goal was to determine an emotionally neutral color palette that we could use in further experiments to isolate the emotional impact of other aspects of visualizations. The second goal was to replicate the results from the study on affective colors [7] with the SAM rating scale.

4.1 S1 – Method

Design. This study had five conditions (the five color palettes). Four come from Bartram et al. [7]'s work on affective color in visualization: the exciting (P_{exc}), positive (P_{pos}), negative (P_{neg}) and calm (P_{cal}) palettes, and we added a grey color palette (P_{gre}). With three repetitions, this study consisted of six blocks (three alternating arousal and valence blocks). Each block included one instance of each of the five palettes.
Datasets. Like the four color palettes from previous work [7], our grey-scale palette has five values, set by linear interpolation between 0 and 1 on the lightness dimension of the CIELAB color space: $(0.166,0,0), (0.333,0,0), (0.5,0,0), (0.666,0,0)$, and $(0.833,0,0)$. Figure 4 shows the color palettes used in this study.
Participants. 50 participants participated. They took on average 8 minutes 9 seconds to complete the study and were compensated £1.5.

4.2 S1 – Hypotheses

We formulated two hypotheses for S1:

H1: P_{cal}, P_{exc}, P_{neg} and P_{pos} will induce the same emotion as in Bartram et al.'s work [7]: P_{cal} will be low and P_{cal} high; P_{exc} and P_{exc} high; P_{neg} and P_{neg} low; and P_{pos} and P_{pos} high.
H2: P_{gre} and P_{gre} will have the most neutral values of the five palettes because the other palettes were designed to induce emotion via color and P_{gre} can be seen as having no color.

4.3 S1 – Results

No responses were excluded. In total, we gathered 1500 trial answers (50 participants × 5 color palettes × 2 measurements × 3 repetitions).

To answer **H1** and **H2** we look at Figure 5. **H1** is confirmed, with P_{cal} being low and P_{cal} being high; both P_{exc} and P_{exc} being high; P_{neg} being low and P_{neg} being neutral to low; and both P_{pos} and P_{pos} being high. **H2** is rejected, with both P_{gre} and P_{gre} being low.

4.4 S1 – Discussion

Confirming **H1** (our results are consistent with those from Bartram et al. [7]) tells us that we can use the SAM scales to capture emotions and that our results can be used as benchmarks for future studies.

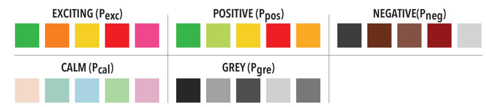

Fig. 4: Color palettes used in S1.

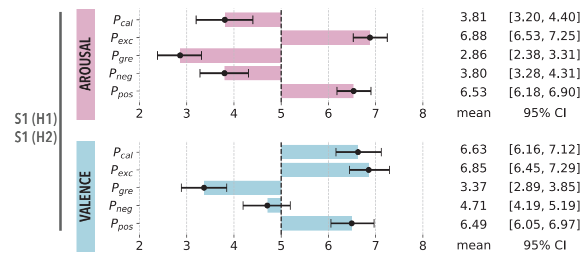

Fig. 5: Mean results for each color palette for S1, centered on a neutral value of 5.

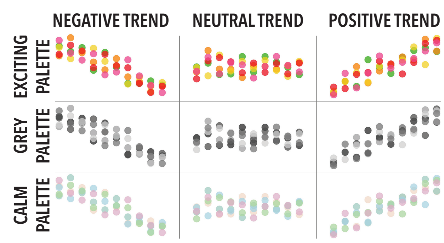

Fig. 6: One set of charts for all 9 conditions used in S2 (three data trends and three color palettes).

Although we did not establish a neutral palette (**H2** rejected), P_{gre} is an appropriate baseline for studying the amount of emotion induced by other factors, given it has the lowest effect on emotion overall.

5 S2 – Color Palettes and Trends

S1 showed color palettes outside a data visualization context, but data visualizations that show real datasets exhibit some characteristics. One such characteristic is the data trend, which can give the impression of motion, which we know affects emotions (e. g., [6, 41]). In this study, we look at the effect of both color and trend on emotions.

5.1 S2 – Method

Design. We used the most extreme palette in each quadrant because they are the most representative of a certain emotion: the exciting (P_{exc} high, P_{exc} high), calm (P_{cal} low, P_{cal} high), and grey (P_{gre} low, P_{gre} low) palettes. We only used three palettes because S1 did not reveal any color palette having high arousal and low valence. We combined them with three trends: positive (T_{pos}), neutral (T_{neu}), and negative (T_{neg}). Three color palettes and three trends resulted in nine conditions. With two repeated measures for each condition, there were two arousal blocks and two valence blocks. Each block had nine graphs – one for each of the nine conditions. We used a different dataset for each repetition, generated as explained in the next section. We used scatterplots to represent the datasets because they are commonly used to represent data trend, with a balanced sampling of each color from a given palette.

Datasets. We generated the experimental datasets with negative, neutral, and positive trends using the following procedure:

1. Define a desired correlation, ρ, and define $\theta = \arccos(\rho)$;
2. Define a vector x_1 for x-coordinates and a random vector x_2 sampled from $\mathcal{N}(2, 0.5)$ of the same length;
3. Center x_1 and x_2 so that they have mean 0, yielding $\dot{x}_1$ and $\dot{x}_2$;
4. Make $\dot{x}_2$ orthogonal to $\dot{x}_1$ by projecting it onto the orthogonal subspace defined by $\dot{x}_1$, yielding $\dot{x}_2^{\perp}$;
5. Scale $\dot{x}_1$ and $\dot{x}_2^{\perp}$ to length 1, yielding $\bar{x}_1$ and $\bar{x}_2^{\perp}$;
6. Lastly define a vector $y = \bar{x}_2^{\perp} + \frac{1}{\tan\theta} \cdot \bar{x}_1$;
7. x_1 and y now have the desired correlation ρ.

We set ρ values of -0.9 for T_{neg}, 0 for T_{neu}, and 0.9 for T_{pos}. We did not choose ρ values of -1 and 1 so that the plots exhibit some realism. To each x-coordinate corresponds five data points – one for each color of the palette. Figure 6 shows one set of graphs used in this study.

Procedure. In addition to the procedure described in Section 3, at the end of each block, participants rated whether their responses were affected more by the colors or the trend of the graph, on a 5-point Likert-scale ranging from "mainly color" to "mainly trend".

Participants. 50 participants participated. They took on average 7 minutes 11 seconds to complete the study and were compensated £1.2.

5.2 S2 – Hypotheses

We formulated four hypotheses for S2:

H1: arousal and valence for each color palette will be consistent with results from S1, as the data has no meaning and we have controlled for trend by averaging over negative, neutral, and positive trends.

H2: $T_{pos} > T_{neu}$ and $T_{neu} > T_{neg}$ because positive trends generally represent good things while negative trends generally represent bad things. Further, Bartram and Nakatani [6] found that abstract motions in the upper-side of a given space are rated positively.

H3: arousal will not be affected by trend because we did not gather indications from previous research of such an effect.

H4: arousal and valence will be more affected by color (i.e. will lead to more variance) than by trend, as there is a well-established, clear effect of color on emotion.

5.3 S2 – Results

We discarded the data from one participant who provided the same answers for all trials. In total, we gathered 1750 trial answers (49 participants $\times$ 3 color palettes $\times$ 3 trends $\times$ 2 measurements $\times$ 2 repetitions) and 196 Likert-scale ratings (49 participants $\times$ 4 blocks).

Effect of color. To answer **H1**, we examine pairwise differences between the ratings of palettes from S1 and from S2, regardless of trend (see Figure 7). The CI for P_{exc} does not overlap with 0, which is strong evidence that P_{exc} is greater in S1 than in S2. On the other hand, the CI for P_{exc} slightly overlaps with 0, so there is no evidence of such difference for P_{exc}. P_{cal} is similar in both studies, but there is strong evidence that P_{cal} is greater in S1 than in S2. Both P_{gre} and P_{gre} are similar in both studies. Figure 8 summarizes these results.

Effect of trend. To answer **H2** and **H3**, for each measure, we compute and bootstrap the signed difference between each pair of trends (see Figure 9). There is strong evidence that $T_{pos} > T_{neu}$ but the effect is small given the measurable range from -8 to $+8$. This effect is even smaller for $T_{neu} > T_{neg}$. There is similar evidence that $T_{pos} > T_{neu}$ and $T_{neu} > T_{neg}$. By transitivity, $T_{pos} > T_{neu} > T_{neg}$ and $T_{pos} > T_{neu} > T_{neg}$.

Effect of color and trend. To answer **H4**, we analyze the differences between the standard deviations σ for color and trend (the larger σ, the larger the effect). We first computed $\sigma_{color} = \{G_{exc}, G_{cal}, G_{gre}\}$, $\sigma_{color} = \{G_{exc}, G_{cal}, G_{gre}\}$, $\sigma_{trend} = \{T_{pos}, T_{neu}, T_{neg}\}$, and $\sigma_{trend} = \{T_{pos}, T_{neu}, T_{neg}\}$ for each participant. Then, we bootstrapped the differences $\sigma_{color} - \sigma_{trend}$ and $\sigma_{color} - \sigma_{trend}$ for each participant (see Figure 10). There is strong evidence that $\sigma_{color} > \sigma_{trend}$ and $\sigma_{color} > \sigma_{trend}$, and the differences are quite large.

Self-reported perceptions. Figure 11 shows that over two-thirds of the participants found that color affected both measures more than trend.

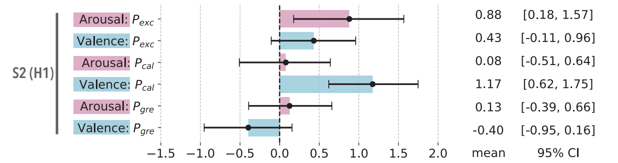

Fig. 7: Pairwise differences between results from S1 and S2.

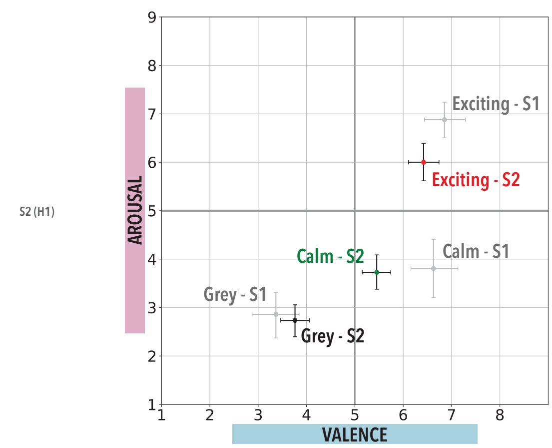

Fig. 8: Mean arousal and valence for color palettes from S2 (opaque color, suffixed with -S2) and S1 (light grey, suffixed with -S1).

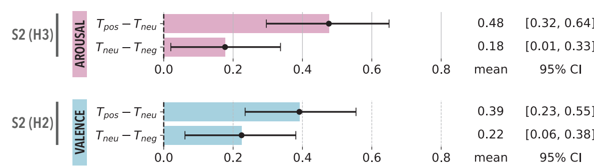

Fig. 9: S2: Differences between trends (all color palettes).

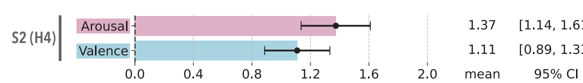

Fig. 10: S2: Pairwise difference between σ_{Color} and σ_{Trend}.

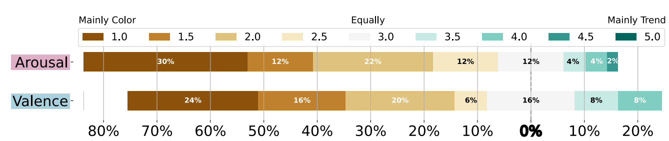

Fig. 11: S2: Participants' self-reported ratings on whether trends or color affect their choices more. Percentages are rounded.

5.4 S2 – Discussion

H1 is only partially confirmed because we found differences between the results from S1 and S2, namely for P_{exc} and for P_{cal}. However, as all color palettes remain in their same quadrant in the arousal-valence chart (see Figure 8), the conveyed emotion of each palette remains consistent. We thus conclude that although there are small variations with and without data, color palettes can be studied without data visualizations and still provide meaningful results. This also makes an important link between the study of color in psychology and in visualization, as the presentation of color stimuli in the two fields generally differs [7, 70].

Regarding the effect of trend in isolation, the results support **H2** but not **H3**. Yet, we interpret this result with caution because the effect sizes are relatively small compared to the effects of color.

The results strongly support **H4** for both measurements. Our discussion of **H2** and **H3** as well as participants' self-reported perceptions demonstrate that color has a greater impact on emotions than trend.

In addition to our confirmatory analysis, through exploratory analysis, we found $T_{pos} - T_{neu} > T_{neu} - T_{neg}$ and $T_{pos} - T_{neu} > T_{neu} - T_{neg}$ for all the palettes we tested. This suggests that the effect of trend on emotion is not linear and that this effect should be particularly considered in visualizations that exhibit strong positive data trends.

6 S3 – Chart Types and Data Trends

We conducted S2 with scatterplots for their common usage and ability to represent trends. As we know, chart types influence people's perception of shapes and trends in data, which in turn affects people's emotions. In S3, we study the effects of chart type and data trend on emotions.

6.1 S3 – Method

Design. We determined which charts to study so that they would: 1) be widely used in visualization; 2) be well-suited to conveying data trends; and 3) support adjusting the amount of ink that is being used to represent data, so that we can mitigate the potential confound that color can create, by using the same amount of ink for all charts. Using these criteria, we selected four types of charts: bar chart (C_{bar}), scatterplot (C_{sca}), jagged line chart (C_{lin}), and smoothed line chart (C_{sml}). We included two types of line charts because this allows us to study in a visualization context the well-known differences in emotional response to jagged vs smooth objects [2,4]. We kept the three trends used in S2: T_{pos}, T_{neu}, and T_{neg}. Four types of charts and three trends resulted in twelve conditions. With three repeated measures for each condition, there were three arousal blocks and three valence blocks. Each block had twelve graphs – one for each of the twelve conditions. We used a different dataset for each repetition, generated as explained next.

Datasets. We generated the datasets with different trends as explained in Section 5.1. To control the amount of ink used in each chart, we varied the visual elements of the charts, calculated the percent of colored pixels for each variation, and selected the one with the desired percentage (6% in this study). Figure 12 shows one set of 12 charts.

Procedure. In addition to the procedure described in Section 3, at the end of the study, participants rated whether their responses were affected more by chart type or by the trend shown in the chart, on a 5-point Likert-scale ranging from "mainly type" to "mainly trend".

Participants. 50 participants participated. They took on average 11 minutes 50 seconds to complete the study and were compensated £1.65.

6.2 S3 – Hypotheses

We formulated six hypotheses for S3:

H1: when aggregating over all chart types, $T_{pos} > T_{neu} > T_{neg}$ and $T_{pos} > T_{neu} > T_{neg}$, following results from S2.

H2: $T_{pos} - T_{neu}$ and $T_{pos} - T_{neu}$ will be greater for C_{lin}, C_{sml} and C_{sca} than for C_{bar} because they make trends more noticeable [16].

H3: $T_{neu} - T_{neg}$ and $T_{neu} - T_{neg}$ will be greater for C_{lin}, C_{sml} and C_{sca} than for C_{bar}, for the same reason.

H4: $C_{lin} > C_{sml}$, as jagged shapes are associated with higher arousal than smooth shapes [2,4].

H5: $C_{lin} < C_{sml}$, as jagged shapes are associated with lower valence than smooth shapes [2,4].

H6: arousal and valence will be more affected by trend than by chart type, as results from S2 show that the trend affects emotion, while there is no known evidence of effects of chart types on emotions.

6.3 S3 – Results

No responses were excluded. In total, we gathered 3600 trial answers (50 participants × 4 chart types × 3 trends × 2 measurements × 3 repetitions) and 100 Likert-scale ratings (50 participants × 2 blocks).

Effect of chart types. We look at the differences between chart types by aggregating all repetitions and trends (see Figure 13). There is clear evidence that C_{bar} is smaller than C_{sml}, C_{lin} and C_{sca}. There is also weak evidence that $C_{sca} < C_{lin}$ and $C_{sca} < C_{sml}$. For valence, all CIs overlap with 0 – there is only inconclusive indication that perhaps $C_{lin} < C_{sml}$, as the CI slightly overlaps with 0.

Effect of trend. To answer **H1**, we compare the effect of trend for all chart types combined (see Figure 14). There is strong evidence that

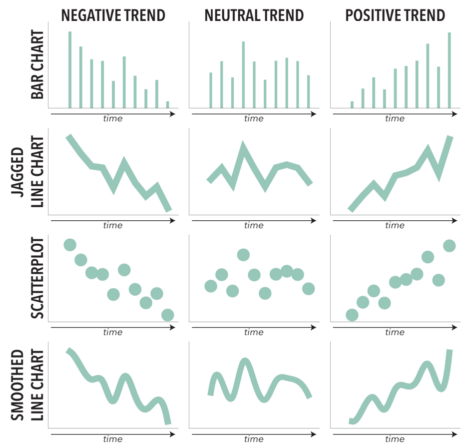

Fig. 12: One set of charts for all 12 conditions used in S3 (three data trends and four chart types), with a "Time" label on the y axis.

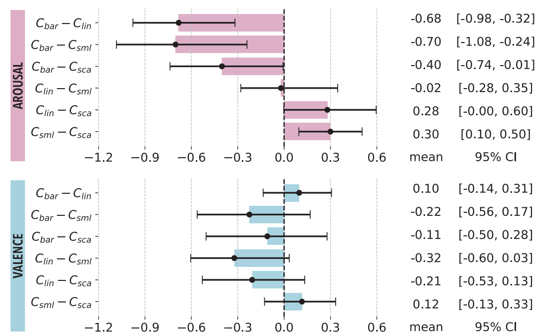

Fig. 13: S3: Differences between chart types (all trends).

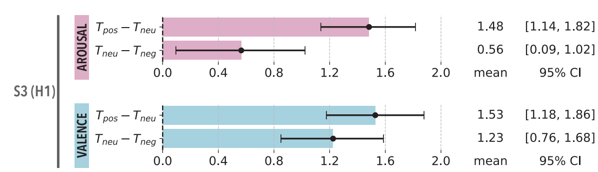

Fig. 14: S3: Differences between trends (all chart types).

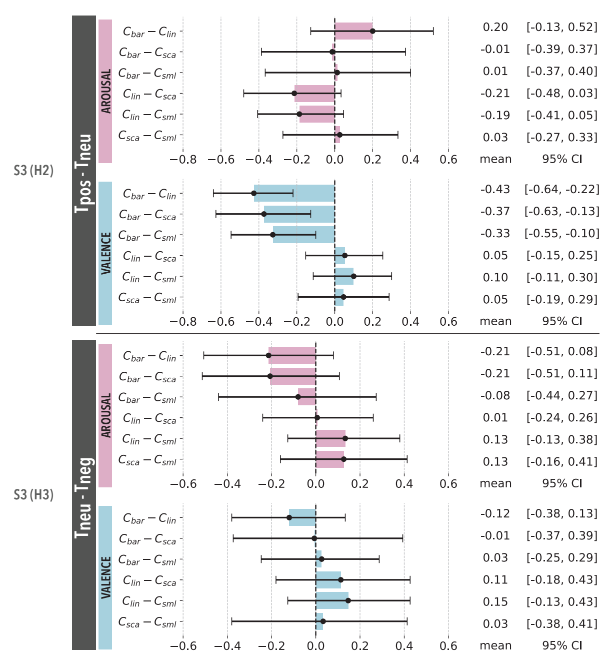

Fig. 15: S3: $T_{pos} - T_{neu}$ and $T_{neu} - T_{neg}$ differences between chart types.

$T_{pos} > T_{neu}$ with a relatively large effect, and that $T_{neu} > T_{neg}$, but with a smaller effect. There is strong evidence that $T_{pos} > T_{neu}$ and that $T_{neu} > T_{neg}$, both with relatively large effects. In sum, there is clear evidence that $T_{pos} > T_{neu} > T_{neg}$ and $T_{pos} > T_{neu} > T_{neg}$ by transitivity.

Difference between T_{pos} and T_{neu} for different chart types. To answer **H2**, we compute the difference $T_{pos} - T_{neu}$ for each chart type and for each measurement after aggregating the repetitions. Then, we calculate and bootstrap the pairwise differences between chart types (see Figure 15). For arousal, the notable differences are that $C_{lin} > C_{sca}$ and $C_{lin} > C_{sml}$; the evidence is weak and the effects are small. For valence, the notable differences are that $C_{bar} > C_{lin}$, $C_{bar} > C_{sca}$ and $C_{bar} > C_{sml}$; the evidence is relatively strong and the effects are small.

Difference between T_{neu} and T_{neg} for different chart types. To answer **H3**, we analyze the differences $T_{neu} - T_{neg}$ and $T_{neu} - T_{neg}$ for each chart type using the same method as for **H2** (see Figure 15). For arousal, the notable differences are that $C_{lin} > C_{sca}$ and $C_{lin} > C_{sml}$; the evidence is weak and the effects are small. For valence, no difference was found, as all CIs of differences largely overlap with 0.

Difference between jagged and smooth line charts. To answer **H4** and **H5**, we analyze the differences $C_{lin} - C_{sml}$ and $C_{lin} - C_{sml}$ for each trend and for all trends combined (see Figure 16). For arousal (**H4**), no difference was found. For valence (**H5**), there is evidence of a difference for T_{neg}, and some possible but inconclusive differences for T_{neu} and all trends combined because the CIs slightly overlap with 0.

Effect of trend and chart type. To answer **H6**, we use the standard deviation (σ) like in S2 (see Figure 17). There is strong evidence that $\sigma_{trend} > \sigma_{type}$ and $\sigma_{trend} > \sigma_{type}$, and the differences are quite large.

Self-reported perceptions. Participants found that trend had a bigger influence on both measurements than chart type (see Figure 18).

6.4 S3 – Discussion

Results from S3 that $T_{pos} > T_{neu} > T_{neg}$ and $T_{pos} > T_{neu} > T_{neg}$ (**H1** confirmed) allow us to generalize results from S2 to other visualizations than scatterplots. Results also confirmed **H6**, with trend having a stronger effect than chart type on both measures. Together, these results indicate that trend plays an important role in determining the emotion a visualization induces, even with no semantic meaning.

We formulated **H2** and **H3** because trend is more easily observed in line graphs and scatterplots than in bar charts [16], therefore we expected that the effect of trend on emotion would be more pronounced in these charts. However, **H2** is only partially supported, as $T_{pos} - T_{neu}$ is smaller for bar charts and $T_{neu} - T_{neg}$ is not significantly different for any of the chart types. Moreover, **H3** is rejected, as for $T_{pos} - T_{neu}$ and $T_{neu} - T_{neg}$ is not significantly different for any of the chart types. Together, these findings indicate that positive trends have a stronger impact on valence in charts where the trend is easier to observe.

Surprisingly, we found no evidence of $C_{lin} > C_{sml}$ and only weak evidence of $C_{lin} < C_{sml}$, which means we do not replicate existing findings that smooth shapes are preferred to sharp shapes [2, 4, 37]. We speculate this might be due to the data trend having a dominating effect on emotion compared to chart type, as shows our confirmation of **H6**.

7 S4 – Chart Types and Data Variance

Given that we found trend has a strong effect on emotion in S2 and S3, there is reason to believe that other data characteristics could also affect emotion. In this fourth experiment, we study the effects of data variance (another data characteristic) and chart type on emotion.

7.1 S4 – Method

Design. We determined three variance levels – low (V_{low}), medium (V_{med}) and high (V_{high}) – through visual inspection. We started by determining V_{low} and V_{high} so that they would produce graphs that are realistic while exhibiting clearly low variance for V_{low}, and while having a visibly erratic nature for V_{high} (with minimum and maximum values nearing the y-scale's lower and upper bound). We then determined V_{med} so that it would produce graphs with peaks and trough magnitudes between those of the graphs generated with V_{low} and V_{high}. The resulting variance levels are $V_{low} : \sigma^2 = 0.1$ ($\sigma = 0.32$), $V_{med} : \sigma^2 = 1.5$

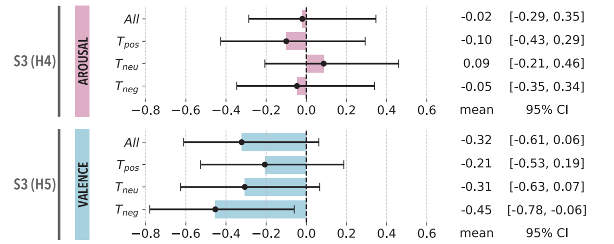

Fig. 16: S3: Differences between chart types of $C_{lin} - C_{sml}$ for each trend.

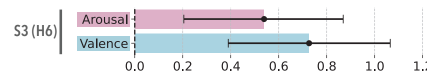

Fig. 17: S3: Pairwise difference between σ_{trend} and σ_{type}.

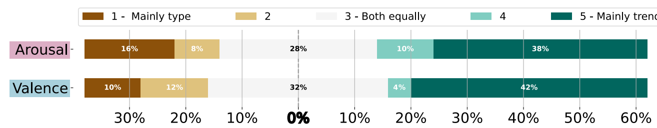

Fig. 18: S3: Participants' self-reported ratings on whether chart type or trend affect their choices more.

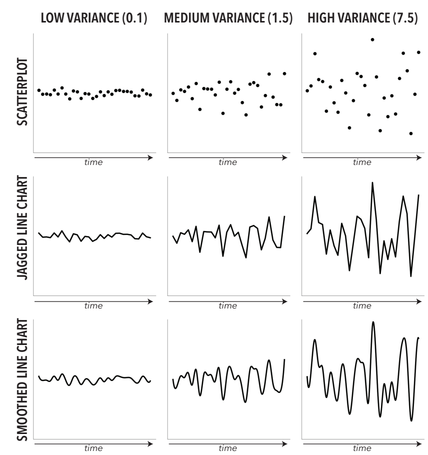

Fig. 19: One set of charts for all 9 conditions used in S4 (three variance levels and three chart types), with a "Time" label on the y axis.

($\sigma = 1.22$) and $V_{high} : \sigma^2 = 7.5$ ($\sigma = 2.74$). To exhibit the variance in each graph, they all have the same y-scale with a range of $(-2, 12)$.

We used a neutral trend ($\rho = 0$) for each graph to isolate the effects of variance and chart type. We studied three chart types: C_{sca}, C_{sml} and C_{lin} (we discarded C_{bar} that is less relevant to studying variance). With three levels of variance and three chart types, there were 9 conditions. With three repeated measures for each condition, there were three arousal blocks and three valence blocks. Each block had nine graphs – one for each of the nine conditions. We used a different dataset for each repetition, and each condition showed the same datasets.

Datasets. We generated the datasets with different variance levels by first defining an array of 30 linearly spaced values ranging from 0 to 10 as the x-values. Then, we defined an array of $y-values$, initially populated with the value 5 30 times. We added variance to each value in the $y-values$ array by drawing from a random normal distribution with a mean of 0 and a standard deviation equivalent to the square root of the target variance. This procedure was reiterated multiple times, and through rejection sampling, datasets of the target variance (± 0.03) and with a correlation coefficient of 0 (± 0.03) were selected. For each variance level, three datasets were generated, corresponding to the three repetitions. Figure 19 shows one set of 9 charts.

Participants. 50 participants participated. They took on average 11 minutes 26 seconds to complete the study and were compensated £1.6.

7.2 S4 – Hypotheses

We formulated five hypotheses for S4:

H1: when aggregating over all chart types, $V_{high} > V_{med} > V_{low}$ because uncertainty is known to induce anxiety [21], which is associated with high arousal [5].

H2: when aggregating over all chart types, $V_{high} < V_{med} < V_{low}$ because uncertainty is known to induce anxiety [21], which is associated with low valence [5].

H3: $C_{lin} > C_{sml}$, as jagged shapes are associated with higher arousal than smooth shapes [2, 4].

H4: $C_{lin} < C_{sml}$, as jagged shapes are associated with lower valence than smooth shapes [4] and S3 provided some evidence of this.

H5: arousal will be higher for line charts than for scatterplot based on S3 results, with $C_{sml} > C_{sca}$ (**H5a**) and $C_{lin} > C_{sca}$ (**H5b**).

7.3 S4 – Results

No responses were excluded. We gathered 2700 trial answers (50 participants $\times$ 3 chart types $\times$ 3 levels of variance $\times$ 2 measurements $\times$ 3 repetitions).

Effect of variance. To answer **H1** and **H2**, we compare the effect of variance for all chart types combined (see Figure 20). There is strong evidence that $V_{high} > V_{med}$ and that $V_{med} > V_{low}$, the latter with a larger effect (**H1** confirmed). There is also evidence that $V_{high} < V_{med}$ but no difference was found between V_{med} and V_{low} (**H2** partially confirmed).

Effect of smoothing. To answer **H3** and **H4**, we analyze $C_{sml} - C_{lin}$ and $C_{sml} - C_{lin}$ (see Figure 21). We found no evidence that $C_{lin} > C_{sml}$ (**H3** rejected), but evidence that $C_{lin} < C_{sml}$ (**H4** confirmed).

Effect of chart type. To answer **H5**, we analyze $C_{sml} - C_{sca}$ and $C_{lin} - C_{sca}$ (see Figure 21). There is strong evidence that $C_{sml} > C_{sca}$ (**H5a** confirmed) and $C_{lin} > C_{sca}$ (**H5b** confirmed).

7.4 S4 – Discussion

Results from S4 support **H1** and replicate previous findings from psychology that unpredictability induces anxiety [21], which is in turn associated with high arousal [5]. Interestingly, the difference is larger for $V_{med} - V_{low}$ than for $V_{high} - V_{med}$. This suggests an effect with diminishing returns wherein arousal increases more quickly as variance increases from low to medium levels than from medium to high levels.

Our results only partially confirm **H2**. Given the relationship between unpredictability and anxiety (which is associated with low valence [5, 21]), our result that $V_{high} < V_{med}$ was expected. However, we also expected to find $V_{med} < V_{low}$. This result, contrasted with the large difference for $V_{med} - V_{low}$, provides further evidence to support the independence of the arousal and valence dimensions of emotion.

Like in S3, we found no evidence of $C_{lin} > C_{sml}$ (**H3** rejected), which means we do not replicate findings that smooth shapes are preferred to sharp shapes [2, 4, 37]. We speculate this might be because the strong effect of variance hides the effect of smoothing on arousal. On the other hand, we did find that $C_{sml} > C_{lin}$ (**H4** confirmed), and this result, combined with S3's result that $C_{sml} > C_{lin}$ for negative trends, shows that smoothing can be used to increase the valence of line charts.

Like in S3, we found strong evidence that both types of line charts we tested have higher arousal than scatterplots, confirming **H5**.

8 S5 – Chart Types and Data Density

Given that both data trend and variance strongly impact arousal and valence, we extend our investigation to study the effects of data density (another data characteristic) and chart type on emotion.

8.1 S5 – Method

Design. Similar to S4, we determined three density levels – low (D_{low}), medium (D_{med}) and high (D_{high}) – through visual inspection. We started by determining D_{low} so that it would produce graphs that are realistic but not too simple. Then, we selected D_{high} so that it would produce graphs that display as much data as possible while their peaks and troughs are still visually separable. We then determined D_{med} so that it would produce graphs with a number of peaks and troughs between that of D_{low} and D_{high}. The resulting density levels are D_{low}: 10 points, D_{med}: 20 points, and D_{high}: 100 points. All datasets have a mean of 5 and all graphs have the same y-scale with a range of $(-2, 12)$.

To isolate the effects of density and chart type, we used a neutral trend ($\rho = 0$) for each graph. We again studied C_{sca}, C_{sml} and C_{lin} because these chart types are commonly used in scenarios with varying degrees of data density. With 3 levels of density and 3 chart types, there were 9 conditions. With 3 repeated measures for each condition, there were three arousal blocks and three valence blocks. Each block had 9 graphs – one for each of the nine conditions. We used a different dataset for each repetition, and each condition showed the same datasets.

Datasets. We generated datasets with different density levels by first defining arrays of 10, 20 and 100 linearly spaced values on the x-axis. Then, we drew a y value from a random normal distribution ($\mu = 5$, $\sigma = 1$) for each x value. We repeated this process to select datasets that had similar variance (± 0.03) and a correlation coefficient near 0 (± 0.03). For each density level, three datasets were generated, corresponding to the three repetitions. Figure 22 shows one set of 9 charts.

Participants. 50 participants participated. They took on average 10 minutes 19 seconds to complete the study and were compensated £1.6.

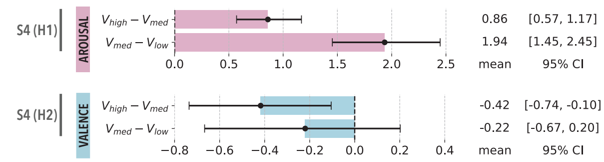

Fig. 20: S4: Differences between variance levels (all chart types).

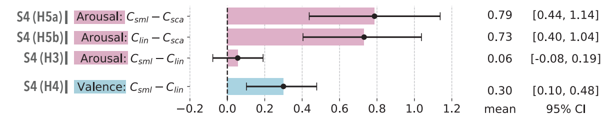

Fig. 21: S4: Differences between chart types (all variance levels).

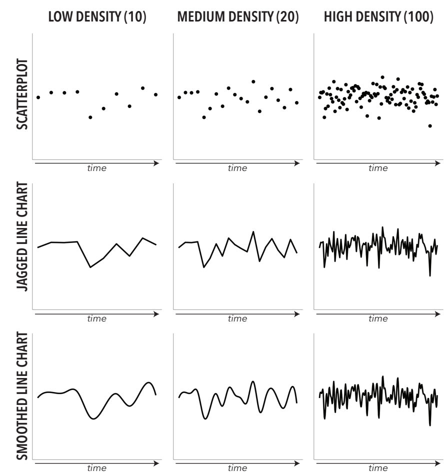

Fig. 22: One set of charts for all 9 conditions used in S5 (three density levels and three chart types), with a "Time" label on the y axis.

8.2 S5 – Hypotheses

We formulated six hypotheses for S5:

H1: when aggregating over all chart types, $D_{high} > D_{med} > D_{low}$ because both increased cognitive load [45] and increased task complexity [46], that density should affect, result in increased arousal.

H2: for C_{sca}, $D_{low} < D_{med} < D_{high}$ because we know that uncertainty is associated with low valence [21] and higher density should make the flat trend more discernible and less uncertain [36].

H3: for C_{sml} and C_{lin}, $D_{low} > D_{med} > D_{high}$ because uncertainty is associated with low valence [21] and higher density should make the charts appear more chaotic and more uncertain.

H4: $C_{sml} < C_{lin}$, like in S3 and S4.

H5: $C_{sml} > C_{lin}$, like in S3 and S4.

H6: arousal will be higher for line charts than for scatterplot like in S4, with $C_{sml} > C_{sca}$ (**H6a**) and $C_{lin} > C_{sca}$ (**H6b**).

8.3 S5 – Results

No responses were excluded. In total, we gathered 2700 trial answers (50 participants × 3 chart types × 3 levels of density × 2 measurements × 3 repetitions).

Effect of Density. To answer **H1**, we compare the effect of density on arousal for all chart types combined (see Figure 23). There is evidence that $D_{high} > D_{med}$ and $D_{med} > D_{low}$, the latter with a smaller effect (**H1** confirmed). To answer **H2** and **H3**, we look at each chart type (see Figure 24), and for all find evidence that $D_{high} > D_{med}$ and $D_{med} > D_{low}$, the latter with smaller effects (**H2** rejected, **H3** confirmed).

Effect of Smoothing. To answer **H4** and **H5**, we analyze $C_{sml} - C_{lin}$ and $C_{sml} - C_{lin}$ (see Figure 25). There is some evidence that $C_{sml} < C_{lin}$ but with a small effect and a confidence bound close to 0 (**H4** weakly confirmed). There is strong evidence that $C_{sml} > C_{lin}$ (**H5** confirmed).

Effect of Chart Type. To answer **H6**, we analyze $C_{sml} - C_{sca}$ and $C_{lin} - C_{sca}$ (see Figure 25). There is evidence that $C_{sml} > C_{sca}$ (**H6a** weakly confirmed) and $C_{lin} > C_{sca}$ (**H6b** confirmed).

8.4 S5 – Discussion

Results from S5 support **H1** (higher density leads to higher arousal) and replicate previous findings that high cognitive load, task complexity, and uncertainty are associated with high arousal [21, 45, 46] as increasing data density increases the amount of information to process.

We hypothesized that increasing density in scatterplots would increase valence because humans are averse to uncertainty [21], and a more obvious flat trend [36] would induce less uncertainty. However, our results rejected **H2**. Perhaps the absence of meaning in the data or of visualization task to perform led to low participants' interest in understanding the data or in estimating data characteristics, making uncertainty and its emotional associations less relevant. It might also be that the relationship between varying levels of density and perceived uncertainty in scatterplots is not that clear – this calls for further research. On the other hand, our hypothesis that line charts with high density induce low valence due to their chaotic nature (**H3**) was confirmed.

Results from S5 complement those from S3 and S4 regarding the Kiki-Bouba effect (that rounded stimuli have lower arousal and higher valence than sharp stimuli [2,4]). While in S3 and S4 we did not find evidence that $C_{sml} < C_{lin}$, we did find such evidence in S5 and confirmed **H4**. Similarly, we had found some evidence that $C_{sml} > C_{lin}$ (only with a negative data trend in S3, and when aggregating all variance levels in S4), and we did find further evidence of this in S5 and confirmed **H5**.

Last, like in S3 and S4, we found that both types of line charts we tested have higher arousal than scatterplots, confirming **H6**.

9 What this Means for Visualization Design

These five studies provide the first systematic exploration of the effects of several immutable data characteristics and visual features on emotion. All the factors we tested do affect emotion to a certain degree, even though data had no semantic meaning. Using Figure 1, we summarize our findings and discuss implications for considering emotion when designing visualizations before discussing limitations and future work.

9.1 Summary of findings

Reflection on Method. Reflecting on the method of using SAM and crowdsourcing experiments, the results are consistent between different studies and with previous work (e. g., [7]). This helps us confirm the validity of the study method and the meaningfulness of the results.

Effect of color. Our results replicate findings that the choice of color affects emotion [7]; they further show that the effect of color is the stronger of the ones we studied: color has a stronger effect than trend, which in turn has a stronger effect than chart type. Although the grey palette is not emotionally neutral, this palette elicited the lowest arousal and valence among all conditions in S1 and S2 and exhibited the smallest variation between S1 and S2. Since the grey palette is both stable and on the low end of the arousal/valence scale, it can be used as a baseline to measure the effect of other color palettes on emotion.

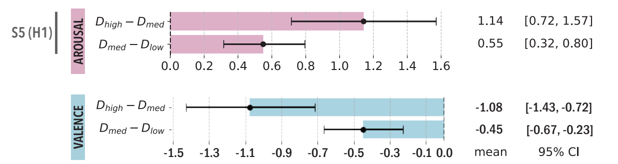

Fig. 23: S5: Differences between density levels (all chart types).

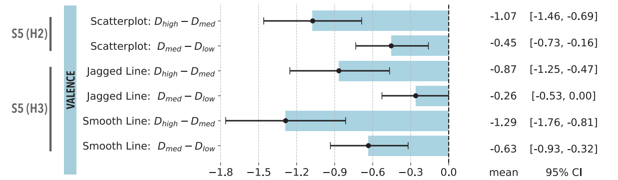

Fig. 24: S5: Differences in valence between density levels for each chart type.

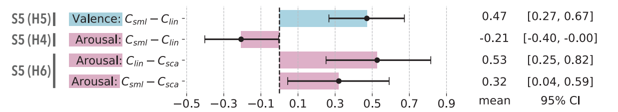

Fig. 25: S5: Differences between chart types (all density levels).

Effect of data characteristics. *Data trend* unsurprisingly has an effect on emotion: positive trends induce higher valence and arousal than neutral trends, that in turn induce higher valence and arousal than negative trends (from S2 and S3). *Data variance* also has an effect: higher variance leads to higher arousal and lower valence; in particular, arousal ratings are more sensitive at the low end of the variance scale, with diminishing sensitivity as variance increases. *Data density* also has an effect: higher density leads to higher arousal and lower valence. This means that data trend, variance and density should be taken into account when designing visualizations aimed at conveying certain emotions.

Effect of chart type. We did not formulate hypotheses on the ordering of chart types in terms of induced arousal and valence due to the lack of previous research on that topic. However, exploratory analysis revealed that line charts (both smooth and jagged) induce higher arousal than scatterplots, which in turn induce higher arousal than bar charts. We also found no evidence that jagged line charts, bar charts and scatterplots would induce different valence; and that the impact of a positive trend on valence (relative to a neutral trend) is higher in chart types which better show trends. This suggests that chart type can affect the perceived valence when the data has a positive trend. In addition, we found some evidence that smoothed line charts induce higher valence than jagged line charts, which corresponds to the known preference for curved objects over sharp ones [4]. Put together, these findings indicate that chart type has an impact on the emotional response of viewers, both independently and in conjunction with the data trend.

9.2 Information Visualization: Emotion for Design

We have found that data trend, data variance, data density, color palette and chart type all have a certain effect on emotion. Below we discuss what these effects mean for visualization design and provide *practical suggestions for designing visualizations while considering emotion.*

Considering Immutable Data Characteristics. Designers have little control over data characteristics such as trend, variance, and density, that do impact emotions. However, designers can change the emotional impact of these characteristics through visualization design choices.

For example, altering the way a trend is perceived will change its impact on arousal and valence. Designers can *employ logarithmic scales or adjust axis start and end values to amplify or diminish trend perception.* For instance, shortening the range of the vertical axis in a visualization with a positive data trend would diminish the perception of that trend and counterbalance the heightened valence and arousal induced by a positive trend. Similar techniques can be applied to data variance and data density. For instance, designers can *increase the range of values of the y axis to make that data appear as if it has lower* *variance* – leading to lower arousal and higher valence. Since high data density is associated with high arousal and low valence, designers could also *make a visualization larger to make the data appear as if it is less dense* – leading to lower arousal and higher valence.

Considering Chart Type. Many guidelines and recommendations exist to choose a chart type, informed by various factors other than emotion (e. g., data types, audience, tasks to support, or implementation constraints). However, our results provide evidence that emotion is yet another aspect that designers should consider when selecting a type of chart, as it can influence the induced emotion, and it can be used to alter the emotional impact of the data trend. For instance, a designer may want to *avoid bar charts if heightening induced valence is a goal.* That is because we found a less pronounced difference in valence ratings for positive and neutral trends in bar charts compared to other chart types. This observation echoes findings that trend is less discernible in bar charts [16], as a greater emotional impact from data trend would be expected when trends are more easily perceived. We also observed that bar charts evoke lower arousal than all other tested chart types, while line charts produce higher arousal than all other tested chart types. As a result, a designer should *prefer line graphs over scatterplots and scatterplots over bar charts if heightening induced arousal is a goal.*

S3, S4 and S5 compared smoothed and jagged line charts, thus providing data about the kiki-bouba effect that is characterized by a preference for rounded objects over sharp ones. We found that smoothed line charts (with round features) induce higher valence than jagged line charts (with sharp features), especially when there is a negative data trend. This means that designers can *smoothen a line chart to increase its induced valence without altering its induced arousal*.

Considering Color. Choosing an appropriate color palette might be the best way for designers to induce emotions. Color is the most impactful factor on emotion among the ones we tested. This finding from the perspective of emotion corroborates previous work that arrived at similar conclusions from the perspective of perception. In Bertin's own (translated) words, *"[color] will always remain the privileged variable of the graphic designer in problems involving differentiation [as color has] an undeniable psychological attraction"* [9, p. 91]. This claim has since then been demonstrated in psychology, where it has been shown that color is more conducive to fast visual search than other encodings such as line orientation [52]; and in visualization, where there is strong evidence of the power of color, in particular for detection tasks [24] and for representing nominal data dimensions [9,43,44]. Our recommendation is, therefore, to *consider color at the end of a visualization design process and as a way to inhibit, balance, or strengthen an emotion based on the data characteristics and based on what other visualization features have been used* – a guideline often suggested based on experience and established visualization guidelines that is given additional importance here, from the perspective of emotion.

Considering the color palette last in visualization design allows designers to consider the emotion induced by other factors (e. g., trend, density, variance, chart type) and to counter or amplify the effect of these factors by using a color palette that induces opposite or similar emotion, respectively. For example, to represent a dataset about an economic crisis with a negative trend in an emotionally neutral way, one could opt for an exciting color palette that would counterbalance the negative emotional response induced by the negative trend.

Possible Side Effects. Making design decisions to impact viewers' emotions does have side effects. For example, positive emotional priming can increase visual judgment performance [23]. Our results suggest that techniques like scaling an axis to exaggerate trends could increase the perceived valence of a visualization and act as a sort of positive visual priming which could lead to improved visual judgment. However, we also know that the distortion induced by the alteration of the axis can significantly alter the perceived message of a visualization [54], and could come at the cost of making comparisons of values more difficult. Sometimes, making such adjustments is difficult. For example, a designer might want to make a visualization with high density larger to increase its induced valence, but they might not have control over the display size. Therefore, designers should consider emotion together with the many other dimensions of visualization design.

9.3 Limitations and Future Work

Like with any controlled experiment, we had to make study design choices. While these choices allow for control and precision, they also result in limitations in terms of generalizability [49].

One limitation of our studies is that the *data trends* were all roughly linear with the same coefficients. We wonder to what extent the form or the intensity of trends could influence the results (e.g., an exponential trend might more strongly affect emotion than a linear trend). Studying trend variations is a promising avenue for future work.

Our results suggest that there might be an interesting non-linear relationship between *data variance* and arousal. Since our experiment was limited to three levels of data variance, we do not have enough granularity to delve deeper into this relationship. Future work could model this relationship to get a clearer picture than we could offer.

In our studies, we only varied the horizontal *data density*. It is possible that vertical data density has a different relationship with emotion than horizontal density. Additionally, we were unable to find a satisfactory explanation for why the valence of scatterplots was inversely related to density. Future work should look into the relationship between data density in scatterplots, correlation perception, and emotion.

We did not consider the accuracy of *information extraction*, which is an essential criterion for designing good visualizations. For example, although the smoothed line charts we used in S3, S4, and S5 display the same data as their non-smoothed version, smoothing creates visual artifacts that are not true to the data. Such modifications (e.g., smoothing, scattering, blurring, aggregating) likely affect the readability of information. Similar concerns apply to the size of a stimulus: the accuracy and efficiency of information retrieval depend on the size of visual idioms; designs that are too small or too large can be hard to read or too cluttered. Future work could look into these aspects by considering both emotional response and task-based performance metrics.

In S3, we controlled for ink ratio, which made the bars in the bar chart thinner than the elements in other chart types. This reduced prominence of the bars may contribute to the lower perceived arousal ratings for bar charts. Future work should compare the effects of ink ratio and feature prominence on induced emotion.

We did not consider the potential impact of culture on emotions. Most study participants were English-speaking from Western countries (see supplemental material), and cultural differences in color meanings [48] and reading direction [29] likely alter visual interpretation. Future work should investigate the role culture plays in this context.

Finally, we excluded *semantics* to isolate the effects of experimental factors, like others have done before (e.g. [7]). However, semantics surely affect a viewer's emotion [30,31], and real-life charts do have semantics. Researchers have argued that data storytelling with visualization could make viewers emotionally connect with the story [20]; and work related to emotion has studied, for example, how storytelling can persuade viewers [53] or make them empathize with a topic [12,38]. Investigating the relationship between semantics and visualization features is promising, especially in a storytelling context.

10 Conclusions

We conducted five studies investigating the relationship between emotion and color, chart type, data trend, data variance, and data density. We found that each has an impact on the emotion induced by a visualization, even if the data has no semantic meaning.

In Study 1, we replicated results from previous work using the SAM rating scales. Then, through Study 2, we quantified the effect of trend on emotion, which we confirmed in Study 3. In Study 3, we also found that chart type has a certain influence on emotion. In Study 4, we investigated the effects of chart type and data variance, confirming previous results about chart type and uncovering the emotional effect of data variance. In Study 5, we tested chart type and data density and quantified the emotional impact of data density.

Put together, these results provide an empirical foundation and practical considerations that designers can use to create charts that account for the emotional impact that visual features of visualization have on viewers. They also pave the way for future research in identifying and quantifying what aspects of a visualization affect emotion.

Acknowledgments

This work was supported by NSERC USRA (#521850) and NSERC Discovery Grant (RGPIN-2019-05422).

11 Supplemental Materials

The supplemental materials, available at `https://osf.io/ywjs4/`, include the following:

- `00-study_material.pdf`: sample pre-block study instructions, and sample valence and arousal stimuli with rating interface.
- `01-demographic-data.pdf`: figure summarizing participant demographics for each study.
- `02-data_column_legend.txt`: description of the columns in the five raw data files.
- `S1-data.csv` to `S5-data.csv`: raw data from each study.
- `S1-analysis.ipynb` to `S5-analysis.ipynb`: analysis script for each study.

References

[1] C. L. Anderson and A. C. Robinson. Affective congruence in visualization design: Influences on reading categorical maps. *IEEE Transactions on Visualization and Computer Graphics*, 28(8):2867–2878, 2022. doi: 10.1109/TVCG.2021.3050118 2

[2] A. Aryani, E. S. Isbilen, and M. H. Christiansen. Affective arousal links sound to meaning. *Psychological science*, 31(8):978–986, 2020. doi: 10.1177/0956797620927967 2, 5, 6, 7, 8

[3] J.-A. Bachorowski. Vocal expression and perception of emotion. *Current directions in psychological science*, 8(2):53–57, 1999. doi: 10.1111/1467-8721.00013 2

[4] M. Bar and M. Neta. Humans prefer curved visual objects. *Psychological science*, 17(8):645–648, 2006. doi: 10.1111/j.1467-9280.2006.01759.x 2, 5, 6, 7, 8

[5] L. F. Barrett. Discrete emotions or dimensions? the role of valence focus and arousal focus. *Cognition & Emotion*, 12(4):579–599, 1998. doi: 10.1080/026999398379574 7

[6] L. Bartram and A. Nakatani. What makes motion meaningful? affective properties of abstract motion. In *2010 Fourth Pacific-Rim Symposium on Image and Video Technology*, pp. 468–474, 2010. doi: 10.1109/PSIVT.2010.85 2, 3, 4

[7] L. Bartram, A. Patra, and M. Stone. Affective color in visualization. In *Proc. CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 1364–1374. ACM, 2017. doi: 10.1145/3025453.3026041 1, 2, 3, 4, 8, 9

[8] Y. Baveye, E. Dellandrea, C. Chamaret, and L. Chen. From crowdsourced rankings to affective ratings. In *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pp. 1–6. IEEE, 2014. doi: 10.1109/ICMEW.2014.6890568 2

[9] J. Bertin. Semiology of graphics: Diagrams, networks, maps, 2011. 9

[10] A. Betella and P. F. Verschure. The affective slider: A digital self-assessment scale for the measurement of human emotions. *PloS one*, 11(2):e0148037, 2016. doi: 10.1371/journal.pone.0148037 2

[11] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE transactions on visualization and computer graphics*, 19(12):2306–2315, 2013. doi: 10.1109/TVCG.2013.234. 2

[12] 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 Conference on Human Factors in Computing Systems*, CHI '17, p. 5462–5474. ACM, 2017. doi: 10.1145/3025453.3025512 9

[13] M. M. Bradley and P. J. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994. doi: 10.1016/0005-7916(94)90063-9 2

[14] M. M. Bradley and P. J. Lang. Measuring emotion: Behavior, feeling, and physiology. In R. D. R. Lane, L. Nadel, G. L. Ahern, J. Allen, and A. W. Kaszniak, eds., *Cognitive Neuroscience of Emotion*, pp. 25–49. Oxford University Press, 2000. 2

[15] A. Cairo. Emotional data visualization: Periscopic's "us gun deaths" and the challenge of uncertainty, 2013. `https://www.peachpit.com/articles/article.aspx?p=2036558`. Accessed: March 2024. 1, 2

[16] M. Correll and J. Heer. Regression by eye: Estimating trends in bivariate visualizations. In *Proc. CHI Conference on Human Factors in Computing Systems*, p. 1387–1396. ACM, 2017. doi: 10.1145/3025453.3025922 5, 6, 9

[17] G. Cumming. The new statistics: Why and how. *Psychological Science*, 25(1):7–29, 2014. PMID: 24220629. doi: 10.1177/0956797613504966 3

[18] C. D'Ignazio and L. F. Klein. Feminist data visualization. In *Workshop on Visualization for the Digital Humanities (VIS4DH)*. IEEE, 2016. 1

[19] P. Dragicevic. *Fair Statistical Communication in HCI*, pp. 291–330. Springer International Publishing, Cham, 2016. doi: 10.1007/978-3-319-26633-6_13 3

[20] N. Gershon and W. Page. What storytelling can do for information visualization. *Commun. ACM*, 44(8):31–37, aug 2001. doi: 10.1145/381641.381653 9

[21] D. W. Grupe and J. B. Nitschke. Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, 14(7):488–501, 2013. doi: 10.1038/nrn3524 7, 8

[22] L. Harrison, R. Chang, and A. Lu. Exploring the impact of emotion on visual judgement. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 227–228. IEEE, 2012. doi: 10.1109/VAST.2012.6400540 1

[23] L. Harrison, D. Skau, S. Franconeri, A. Lu, and R. Chang. Influencing visual judgment through affective priming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, p. 2949–2958. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2470654.2481410 9

[24] C. G. Healey, K. S. Booth, and J. T. Enns. High-speed visual estimation using preattentive processing. *ACM Trans. Comput.-Hum. Interact.*, 3(2):107–135, jun 1996. doi: 10.1145/230562.230563 9

[25] J. Heer and M. Stone. Color naming models for color selection, image editing and palette design. In *Proc. CHI Conference on Human Factors in Computing Systems*, CHI '12, p. 1007–1016. ACM, 2012. doi: 10.1145/2207676.2208547 2

[26] N. Henry Riche. Beyond system logging: human logging for evaluating information visualization. In *In BELIV'10 workshop at ACM SIGCHI 2010*. ACM, April 2010. 2

[27] D. Hermans, J. D. Houwer, and P. Eelen. The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion*, 8(6):515–533, 1994. doi: 10.1080/02699939408408957 1

[28] N. Kaya. Relationship between color and emotion: A study of college students. *College Student Journal*, 38(3):396–405, 2004. 2

[29] S. Kazandjian and S. Chokron. Paying attention to reading direction. *Nature Reviews Neuroscience*, 9(12):965–965, 2008. doi: 10.1038/nrn2456-c1 9

[30] H. Kennedy and R. L. Hill. The feeling of numbers: Emotions in everyday engagements with data and their visualisation. *Sociology*, 52(4):830–848, 2018. doi: 10.1177/0038038516674675 1, 2, 9

[31] X. Lan, Y. Shi, Y. Zhang, and N. Cao. Smile or scowl? looking at infographic design through the affective lens. *IEEE Transactions on Visualization and Computer Graphics*, 27(6):2796–2807, 2021. doi: 10.1109/TVCG.2021.3074582 9

[32] X. Lan, Y. Wu, and N. Cao. Affective visualization design: Leveraging the emotional impact of data. *IEEE Transactions on Visualization and Computer Graphics*, 30(01):1–11, jan 2024. doi: 10.1109/TVCG.2023.3327385 1

[33] 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 1, 2

[34] X. Lan, Y. Wu, Y. Shi, Q. Chen, and N. Cao. Negative emotions, positive outcomes? exploring the communication of negativity in serious data stories. In *Proc. CHI Conference on Human Factors in Computing Systems*, CHI '22. ACM, 2022. doi: 10.1145/3491102.3517530 1

[35] P. J. Lang. Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, and T. A. Williams, eds., *Technology in mental health care delivery systems*, pp. 119 – 137. Norwood, NJ: Ablex, 1980. 1, 2

[36] T. W. Lauer and G. V. Post. Density in scatterplots and the estimation of correlation. *Behaviour & Information Technology*, 8(3):235–244, 1989. doi: 10.1080/01449298908914554 7, 8

[37] H. Leder, P. P. Tinio, and M. Bar. Emotional valence modulates the preference for curved objects. *Perception*, 40(6):649–655, 2011. doi: 10.1068/p6845 2, 6, 7

[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 9

[39] C. Lin and M. A. Thornton. Fooled by beautiful data: Visualization aesthetics bias trust in science, news, and social media. *PsyArxiv*, 2021. doi: 10.31234/osf.io/dnr9s 1, 2

[40] S. Lin, J. Fortuna, C. Kulkarni, M. Stone, and J. Heer. Selecting semantically-resonant colors for data visualization. In *Computer Graphics Forum*, vol. 32, pp. 401–410. Wiley Online Library, 2013. doi: 10.1111/cgf.12127 2

[41] M. Lockyer and L. Bartram. Affective motion textures. *Computers & Graphics*, 36(6):776–790, 2012. 2011 Joint Symposium on Computational Aesthetics (CAe), Non-Photorealistic Animation and Rendering (NPAR), and Sketch-Based Interfaces and Modeling (SBIM). doi: 10.1016/j.cag.2012.04.009 1, 2, 3

[42] K. T. Lyra, S. Isotani, R. C. Reis, L. B. Marques, L. Z. Pedro, P. A. Jaques, and I. I. Bitencourt. Infographics or graphics+ text: Which material is best for robust learning? In *2016 ieee 16th international conference on advanced learning technologies (icalt)*, pp. 366–370. IEEE, 2016. doi: 10.1109/ICALT.2016.83 2

[43] A. M. MacEachren. *How maps work: representation, visualization, and design*. Guilford Publications, 2004. 9

[44] J. Mackinlay. Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, 5(2):110–141, apr 1986. doi: 10.1145/22949.22950 9

[45] M. K. MacPherson, D. Abur, and C. E. Stepp. Acoustic measures of voice and physiologic measures of autonomic arousal during speech as a function of cognitive load. *Journal of Voice*, 31(4):504–e1, 2017. doi: 10.1016/j.jvoice.2016.10.021 7, 8

[46] J. Malmberg, E. Haataja, and S. Järvelä. Exploring the connection between task difficulty, task perceptions, physiological arousal and learning outcomes in collaborative learning situations. *Metacognition and Learning*, 17(3):793–811, 2022. doi: 10.1007/s11409-022-09320-z 7, 8

[47] I. B. Mauss and M. D. Robinson. Measures of emotion: A review. *Cognition and emotion*, 23(2):209–237, 2009. doi: 10.1080/02699930802204677 2

[48] D. McCandess and AlwaysWithHonor.com. Emotional data visualization: Periscopic's "us gun deaths" and the challenge of uncertainty, 2009. https://informationisbeautiful.net/visualizations/colours-in-cultures/. Accessed: June 2024. 9

[49] J. E. McGrath. *Methodology matters: doing research in the behavioral and social sciences*, p. 152–169. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995. 9

[50] A. Mehrabian and J. A. Russell. *An approach to environmental psychology*. the MIT Press, 1974. 2

[51] K. Moreland. Diverging color maps for scientific visualization. In *Advances in Visual Computing*, pp. 92–103. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. doi: 10.1007/978-3-642-10520-3_9 2

[52] H.-C. Nothdurft. The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision research*, 33(14):1937–1958, 1993. doi: 10.1016/0042-6989(93)90020-W 9

[53] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2211–2220, 2014. doi: 10.1109/TVCG.2014.2346419 9

[54] A. V. Pandey, K. Rall, M. L. Satterthwaite, O. Nov, and E. Bertini. How deceptive are deceptive visualizations? an empirical analysis of common distortion techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, p. 1469–1478. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2702123.2702608 9

[55] F. Phillips, J. F. Norman, and A. M. Beers. Fechner's aesthetics revisited. *Seeing and perceiving*, 23 3:263–71, 2010. doi: 10.1163/187847510X516412 2

[56] N. H. Riche, C. Hurter, N. Diakopoulos, and S. Carpendale. *Data-Driven Storytelling*. A. K. Peters, Ltd., USA, 1st ed., 2018. doi: 10.1201/9781315281575 2

[57] B. Rogowitz and L. Treinish. Data visualization: the end of the rainbow. *IEEE Spectrum*, 35(12):52–59, 1998. doi: 10.1109/6.736450 2

[58] J.-P. Royet, D. Zald, R. Versace, N. Costes, F. Lavenne, O. Koenig, and R. Gervais. Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: a positron emission tomography study. *Journal of Neuroscience*, 20(20):7752–7759, 2000. doi: 10.1523/JNEUROSCI.20-20-07752.2000 1

[59] J. A. Russell. Is there universal recognition of emotion from facial expression? a review of the cross-cultural studies. *Psychological bulletin*, 115(1):102, 1994. doi: 10.1037/0033-2909.115.1.102 2

[60] S. Silva, B. S. Santos, and J. Madeira. Using color in visualization: A survey. *Computers & Graphics*, 35(2):320–333, 2011. doi: 10.1016/j.cag.2010.11.015 2

[61] D. H. Silvera, R. A. Josephs, and R. B. Giesler. Bigger is better: the influence of physical size on aesthetic preference judgments. *Journal of Behavioral Decision Making*, 15(3):189–202, 2002. doi: 10.1002/bdm.410 2

[62] D. A. Szafir. Modeling color difference for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):392–401, 2018. doi: 10.1109/TVCG.2017.2744359 2

[63] P. P. Tinio and H. Leder. Just how stable are stable aesthetic features? symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica*, 130(3):241–250, 2009. doi: 10.1016/j.actpsy.2009.01.001 2

[64] P. Valdez and A. Mehrabian. Effects of color on emotions. *Journal of experimental psychology: General*, 123(4):394, 1994. doi: 10.1037/0096-3445.123.4.394 2

[65] A. Vande Moere, M. Tomitsch, C. Wimmer, B. Christoph, and T. Grechenig. Evaluating the effect of style in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2739–2748, 2012. doi: 10.1109/TVCG.2012.221 2

[66] Y. Wang, A. Segal, R. Klatzky, D. F. Keefe, P. Isenberg, J. Hurtienne, E. Hornecker, T. Dwyer, and S. Barrass. An emotional response to the value of visualization. *IEEE computer graphics and applications*, 39(5):8–17, 2019. doi: 10.1109/MCG.2019.2923483 1

[67] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, 3 ed., 2012. doi: 10.1016/C2009-0-62432-6 2

[68] A. B. Warriner, V. Kuperman, and M. Brysbaert. Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior research methods*, 45(4):1191–1207, 2013. doi: 10.3758/s13428-012-0314-x 2

[69] L. Weller and R. Livingston. Effect of color of questionnaire on emotional responses. *The Journal of General Psychology*, 115(4):433–440, 1988. PMID: 3210003. doi: 10.1080/00221309.1988.9710580 2

[70] L. Wilms and D. Oberfeld. Color and emotion: effects of hue, saturation, and brightness. *Psychological research*, 82(5):896–914, 2018. doi: 10.1007/s00426-017-0880-8 2, 4