Seeing Through the Overlap: The Impact of Color and Opacity on Depth Order Perception in Visualization

Zhiyuan Meng
School of Computer Science and Technology
Shandong University
Qingdao, Shandong, China
zhiyuanmeng340@gmail.com

Yunpeng Yang
School of Computer Science and Technology
Shandong University
Qingdao, Shandong, China
kuku718@163.com

Qiong Zeng*†
School of Computer Science and Technology
Shandong University
Qingdao, Shandong, China
qiong.zn@sdu.edu.cn

Kecheng Lu
School of Information
Renmin University of China
Beijing, China
lukecheng0407@gmail.com

Lin Lu
School of Computer Science and Technology
Shandong University
Qingdao, Shandong, China
llu@sdu.edu.cn

Changhe Tu
School of Computer Science and Technology
Shandong University
Qingdao, Shandong, China
chtu@sdu.edu.cn

Fumeng Yang*
Department of Computer Science
University of Maryland College Park
College Park, Maryland, United States
fy@umd.edu

Yunhai Wang
School of Information
Renmin University of China
Beijing, China
wang.yh@ruc.edu.cn

Abstract

Semi-transparent visualizations are commonly used to reveal information in overlapped regions by applying colors and opacity. While a few studies made recommendations on how to choose colors and opacity levels to maintain depth perception, they often conflict and overlook the interaction effect between these factors. In this paper, we systematically explore the impact of color and opacity on depth order perception across eight colors, three opacity levels, and various layer orders and arrangements. Our inferential analysis shows that both color hue and opacity significantly influence depth order perception, with the effectiveness depending on their interaction. We also derived 12 features for predictive analysis, achieving the best mean accuracy of 80.72% and mean F1 score of 87.75%, with opacity assigned to the front layer as the top feature for most models. Finally, we provide a small design tool and four guidelines to better align the design rules of colors and opacity in semi-transparent visualizations.

CCS Concepts

• **Human-centered computing → Visualization design and evaluation methods**.

Keywords

Color Design, Opacity, Depth Order Perception, Semi-Transparent Visualization

1 Introduction

Semi-transparent visualizations often employ color and opacity to represent data with self or mutual occlusion. Their applications include volume renderings [6, 9, 17, 35, 61], scatter plots [36, 39], and parallel coordinates [26, 54] (Fig. 1a). In these visualizations, colors often represent categories or structures, while opacity can encode important temporal information or serve as a redundant visual cue [26]. Appropriately chosen color and opacity can enhance a viewer's ability to understand and analyze data by facilitating depth order perception, helping users identify occluded layers and differentiate between overlapping data items (e.g., samples at different timestamps). In the example cited above, well-chosen color and opacity enable viewers to explore different categories in parallel coordinates without confusion.

However, choosing an appropriate combination of color and opacity is no easy feat. Despite decades of visualization research on color design, studies addressing opacity are surprisingly scarce [62]. This gap is particularly concerning, as humans inherently struggle with depth perception—a challenge further complicated by the complex interplay between multiple layers and varying opacity levels [12, 55]. The limited research leads to contradictory guidelines. For example, while a few studies suggest that increasing the opacity of the front layer enhances depth order perception [1, 54], others argue that opacity has little impact [28, 29].

A crucial oversight in these studies is the neglect of the interaction effects[1] between color and opacity across different layers (see

*Equal mentorship.
†Corresponding author.
[1]Interaction effects and interactive effects are often used interchangeably; we choose interaction effects, which is a statistical term that describes how a third variable influences the relationship between an independent and dependent variable [25].

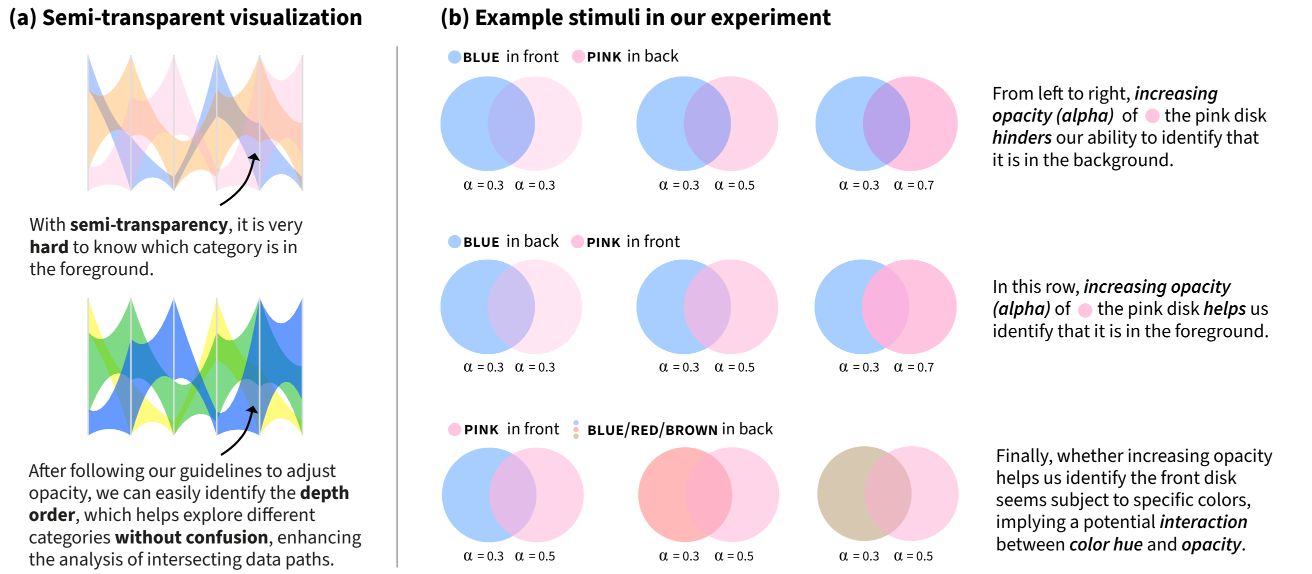

Figure 1: Examples of semi-transparent visualizations. (a) We show examples of semi-transparent parallel coordinates, with the top row displaying a visualization with inadequate semi-transparency, and the bottom row showing its counterparts with appropriate semi-transparency. Additional examples are provided in Appx. A. (b) We show example stimuli in our experiment. We varied color hues and opacity levels in two overlapping disks: one in the background (the back layer) and the other in the foreground (the front layer). We also tested two arrangements: the front disk was positioned either on the left or on the right. Participants were asked to select the disk that appeared to be in the front.

Fig. 1b). These interaction effects likely drive the inconsistencies in the findings, leaving designers and practitioners without clear guidelines.

In this paper, we comprehensively examine the effects of color, opacity, and their interactions on depth order perception, unifying them within a single experiment. While various color attributes (e.g., hue, saturation, lightness) may influence this perception, we focus on color hue—it is the most commonly considered factor in conveying categorical differences and enhancing visual communication [23, 45, 56], and preliminary evidence suggests it has a stronger influence than saturation and lightness [55].

Through inferential and predictive analyses of our data, we contribute **quantitative results** on the effects of color and opacity in visualization (Sec. 4), and **predictive models** for depth order perception that could be integrated into design processes (Sec. 5). Additionally, we conduct exploratory analyses of lightness and saturation on depth perception to complement our findings (Sec. 6). Based on these findings and models, we formulate **a set of actionable design recommendations** (Sec. 7.1) and provide a small design tool to assist visualization designers and practitioners in selecting colors and opacity (see supplementary materials).

Specifically,

- In an online experiment with 192 participants, we tasked participants with selecting the front disk in a pair of overlapping disks (see Fig. 1b) [55]. This experiment examines 1,008 unique combinations of colors, opacity levels, layer orders, and arrangements. Our inferential analysis shows that **both color hue and opacity significantly influence** depth order perception in semi-transparent visualizations, while the interaction effect between color hue and opacity is **mild**. Moreover, *increasing* opacity in the *front* layer, or *decreasing* opacity in the *back* layer *enhances* depth order perception. Among the eight colors tested, BLUE emerges as the most effective color for the front layer, while PINK and YELLOW are the most effective colors for the back layer.
- We also conducted predictive analyses on the experimental data. We derived 12 features (e.g., hue distance) to predict whether viewers can correctly identify the front disk, followed by a sensitivity analysis. Across all our predictive models, except for the SVM with a linear kernel, the top feature for predicting depth order perception was consistently OPACITY FRONT—the opacity assigned to the front layer. Our best-performing model, **a random forest and four features**, achieved **the best average accuracy of 80.72%** and **average F1 score of 87.75%**. These predictive models provide a basis for our small tool, where designers can improve a pair of colors and opacity to get an estimation of viewers' perception of different layers.
- Finally, we offer four design guidelines on color hue and opacity to help designers create effective semi-transparent visualizations. These guidelines enhance visual clarity and depth perception by emphasizing the selection of color hues and the adjustment of opacity based on layer orders.

Our experiment represents the most comprehensive effort to date in evaluating the impacts of color hue and opacity on depth order perception. While our study focuses on a pair of two overlapping disks, our results also provide insights into how overlaps between adjacent layers might behave in more complex scenarios. We anticipate that our results and methodologies will inspire future work examining additional factors such as shape and size, as well as more complex datasets, to fully explore the design space for semi-transparent visualizations. To facilitate these efforts, we made our design tool, data, code, and models available at https://osf.io/n3jg8/.

2 Related Work

Our work relates to semi-transparent visualizations for scientific volume data and abstract information data. We also examine existing color design guidelines tailored to semi-transparent visualizations to help contextualize our work.

2.1 Depth Perception in Semi-Transparent Volumes

Semi-transparent visualization techniques have been extensively explored to visualize scientific volume. An important technique in this area is direct volume rendering (DVR), which incorporates depth signals, such as color, opacity, shading, and lighting, to enhance semi-transparent structures the (see [18, 27, 58] for surveys).

Several perceptual studies highlight the role of depth perception and identify key factors influencing depth perception in semi-transparent volumes. For example, Adelson and Anandan [1] proposed a depth layer representation based on *luminance variations* at X-junctions, which can reveal areas prone to depth order ambiguity. Kersten et al. [28] investigated the interaction between *transparency and motion*, demonstrating that transparency can bias depth perception derived from motion cues. Kersten et al. [29] explored the impact of *opacity* and *spatial frequency* on stereoscopic rendering in absorptive media. They found that opacity has minimal influence on depth perception accuracy. Boucheny et al. [6] employed a three-alternative forced-choice test to evaluate the effectiveness of DVR in static and dynamic scenarios. They emphasized that *transparency* can lead to depth order ambiguities. Englund and Ropinski [17] designed experiments using ordinal and absolute depth judgments to study the impact of six volume rendering techniques on depth perception, shape perception, and visual appeal. Their findings suggest that clear back-to-front relationships encoded through *image contrast* enhance absolute depth perception.

Perceptual studies in this area also offer valuable insights into how humans perceive depth order in semi-transparent visualizations. These insights inspire the development of computational methods to improve depth perception in such data representations. The core idea often involves adjusting transparency based on object distance or leveraging data-driven algorithms. Chan et al. [10] incorporated psychological principles like *visibility*, *shape*, and *transparency* to create quantitative metrics for the perceived quality of layers. Based on these metrics, they developed automatic transparency optimization algorithms that enhance the visualization of layered structures in volume rendering. Zheng et al. [61] proposed a quantitative perceptual model to improve depth order perception using *transparency and luminance*. These studies suggest that cues such as opacity and luminance play a significant role in depth order perception within semi-transparent volume rendering, which subsequently motivates our work.

2.2 Depth Perception in Semi-Transparent Abstract Data

Semi-transparent visualization also plays a key role in information visualization for showing overlapping categorical data [26, 40, 57]. This technique proves particularly useful when data points from distinct categories either share identical values across multiple variables (common in parallel coordinates) or occupy the same spatial location (e.g., scatter plots). For example, Wegman and Luo [57] pioneered the use of *opacity* in parallel coordinates to visualize overlapping lines, significantly improving the readability of large datasets. Johansson et al. [26] built upon this concept, incorporating depth cues to encode temporal information within parallel coordinates. Data in older time steps is drawn in the background with reduced *saturation* and *brightness* using a single hue. Mayorga et al. [37] leveraged semi-transparency to encode multiple attributes, such as group membership and density. They employed *lightness* and *chroma* parameters to emphasize the density of overlapping sets.

Ensuring visually distinct layers is crucial in semi-transparent visualizations for clear comprehension of overlapping categorical data [40]. However, interpreting overlapping visual structures can be challenging. To address this, researchers have conducted empirical studies to understand how various factors associated with semi-transparency influence the perception of depth order, with a particular focus on color and opacity. For instance, Grieco and Roncato [20] demonstrated that both *opacity* and the contour of data *shapes* significantly impact viewers' ability to discriminate between depth orders. Hagh-Shenas et al. [21] identified that blending colors with opponent *hues* can lead to difficulties in visual perception. Bartram et al. [4] investigated the influence of *color and opacity* on overlaid grids. Their findings emphasized the crucial role of color in differentiating layered information. However, existing work also revealed the complex interplay between color and opacity, highlighting the inherent difficulty in predicting their combined effect on depth perception. Our study contributes to this ongoing challenge by unveiling the impact of color and opacity on depth order perception and their interaction effects, with careful consideration of both front and back layers.

2.3 Color Design for Semi-Transparent Visualization

Numerous empirical studies and computational models exist to guide effective color selection within visualizations [8, 33, 46, 51, 60, 62]. However, when applied to semi-transparent visualizations, these established guidelines require significant adaptation due to the interplay between color and opacity in conveying depth order. Unlike opaque visualizations, semi-transparent ones introduce additional layers that interact with colors. This necessitates a systematic investigation of how color and opacity interactively influence depth perception and, consequently, refined color design guidelines tailored for semi-transparent visualization. As a pioneer

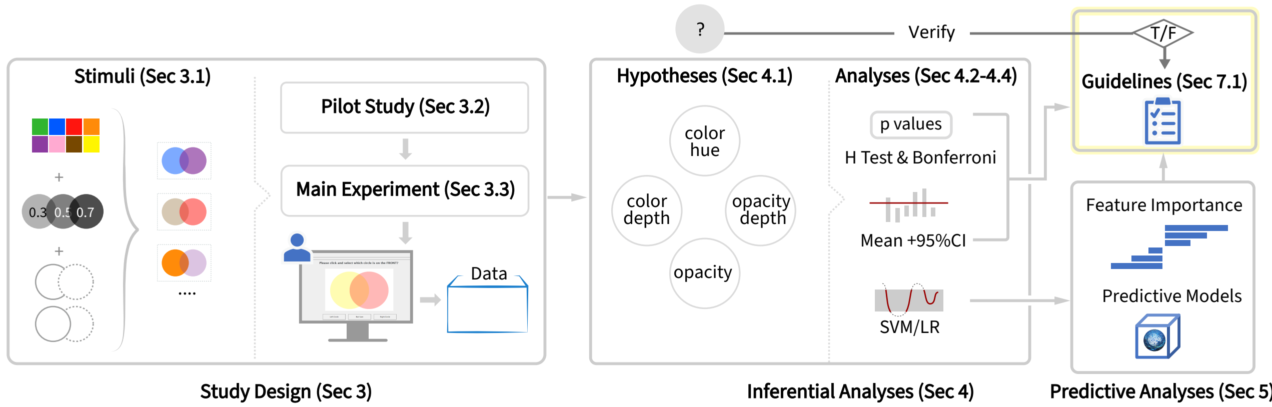

Figure 2: Overview of our research. We begin with the generation of stimuli, followed by a pilot experiment and a main experiment. We conducted *Kruskal-Wallis* H tests with *Bonferroni correction* and calculated bootstrapped confidence intervals. Additionally, we constructed predictive models for depth order perception, which also shows how variations in color hue and opacity between the front and back layers affect the perception.

work, Wang et al. [55] leveraged conjoint analysis with two overlapping disks to explore the contributions of a single parameter value, including color, opacity, lightness, position, and color blending algorithms. They found that color and opacity in the front and back layers are important to convey depth order. Similarly, previous studies have produced a set of limited design considerations for semi-transparent visualizations, which we categorized below:

- **Using colors to enhance depth perception.** Often, it is recommended to use warmer colors (e.g., red or orange) in the foreground and colder colors (e.g., blue or green) in the background, which align with viewers' perception of proximity [3, 52]. However, Wang et al. [55] suggested an contradictory approach to using cold colors in the foreground, overlapped with warm colors in the background. Additionally, increasing the lightness contrast between foreground and background layers can further emphasize depth ordering [12, 17, 55, 56]. For example, using lighter foreground objects against darker backgrounds can increase the perceived depth.
- **Optimizing opacity for depth clarity.** Generally, increasing the opacity of foreground objects can enhance depth ordering accuracy [1, 55]. However, excessively high opacity should be avoided in contexts where transparency conveys critical information or when subtle layering effects are needed [48]. In particular, a few studies suggest that opacity has only a limited influence on the accuracy of depth perception [28, 29].
- **Balancing color and opacity for effective depth perception.** Previous research suggested combining highly saturated colors with lower opacity to enhance visual prominence, as the perception of opacity changes with variations in color saturation and color contrast [10]. Also, placing cold colors in the foreground and warm colors in the background can improve depth perception even at low opacity levels [55].

These considerations reveal two primary limitations of extant studies. First, **they often overlook interaction effects**. Most studies treat color and opacity as separate factors influencing depth perception, whereas others [10, 55] provide limited evidence and are insufficient for providing practicable guidelines to choose appropriate colors and opacity levels. Second, **they produce conflicting design recommendations**. There is a lack of consensus among studies on the appropriate use of color and opacity for depth perception. Motivated by these needs, we followed a consistent and systematic approach to color design in semi-transparent visualizations. We adopted the experimental framework by Wang et al. [55] and examined the design space of color and opacity. Contrasting with previous works, we had a focus on the interplay between these factors and examined a broader range of color and opacity conditions.

3 Study Design

To start, we defined our core research question: **what are the impacts of color, opacity, and their interaction on depth order perception?** This research question guided us throughout the processes of study design and data analysis. Following this research question, we selected eight representative colors based on their hues and generated a set of stimuli to ensure comprehensive coverage. We first conducted a pilot study with 24 participants, which helped us fine-tune our experimental design and determine the appropriate sample size. Subsequently, we conducted our main experiment with a total of 192 crowdsourced participants. Fig. 2 provides an overview of our research framework.

3.1 Stimuli

Paired comparison experiments are extensively used in the fields of visualization [22, 38, 41] and color perception [3, 19, 55]. Any complex comparisons or overlaps can be decomposed into pairs of comparisons between two objects. We adopted this methodology, presenting viewers with two objects. To examine the interaction effect between color and opacity, similar to [55], we generated two equally sized, semi-transparent disks with a 50% overlap in diameter (also see Appx. B). We generated images for all possible permutations of color and opacity between the front and back layers, using the classic α-blending algorithm [43] to calculate a smooth color and opacity transition in the overlapped regions.

Opacity levels. When generating stimuli, we observed that if any disk had opacity below 0.3, depth perception and overlapping identification was overly challenging. However, when a disk's opacity was above 0.7, distinguishing between the two layers became too easy, but making it difficult to identify the overlapping area. As such, we set opacity values at $\alpha = \{0.3, 0.5, 0.7\}$ as low, medium, and high opacity levels.

Color choices. Since color names reflect how people naturally organize and identify colors [23, 56], we first selected frequently-used color names following previous guidelines [5, 55, 56]. Their studies unveil four primary colors (red, green, blue, yellow) and four secondary colors (brown, orange, purple, pink) as most commonly used color terms for categorical tasks, excluding black, white, and gray. We then identified the most representative color values for each specific color name. We used the dataset constructed by Heer and Stone, measured the representativeness as negative entropy, and selected the most representative color value with the minimum negative entropy [23]. This process resulted in a set of eight colors: ● RED, ● ORANGE, ● BROWN, ● YELLOW, ● GREEN, ● BLUE, ● PURPLE, and ● PINK.[2]

Color blending algorithm. We employed the classic α-blending algorithm [43] to simulate the blending effect for the front and back layers as weighted averages:

$$\alpha_{\text{MIX}} = 1 - (1 - \alpha_t)(1 - \alpha_b)$$

$$C_{\text{MIX}} = \frac{C_t \alpha_t + C_b \alpha_b (1 - \alpha_t)}{\alpha_{\text{MIX}}}$$

The notations C_{MIX} and α_{MIX} represent the blended color and opacity within the overlapped region, respectively. The subscripts t and b denote the front and back layers, with C representing color and α representing opacity. By applying various color and opacity combinations to the two disks and positioning them in the foreground or background, we generated a dataset of 1,008 unique visual stimuli.[3] We show examples of our stimuli in Fig. 1b.

3.2 Pilot Study

Task. Building on the work of Wang et al. [55], our study presented participants with two overlapping disks in each trial and asked them to identify **which disk appeared to be in the front**. They could choose "left", "right" and "uncertain" (see Fig. 3). For each participant, we recorded whether each trial was correctly answered (1 for correct, 0 for incorrect) and the time taken to complete it.

[2]The specific HSL (Hue, Saturation, and Lightness) values are ● RED {1,100,52}, ● ORANGE {31,100,51}, ● BROWN {36,100,23}, ● YELLOW {57,100,50}, ● GREEN {119,58,44}, ● BLUE {219,100,50}, ● PURPLE {288,46,42}, and ● PINK {332,100,83}, respectively.

[3]We first select two colors from the eight colors, resulting in $C_8^2 = 28$ combinations, and then assign an opacity value to each color of the two colors, and we have $3 \times 3 = 9$ options. Finally, we determine the layer order and arrangement ($2 \times 2 = 4$ options). So ultimately we have $28 \times 9 \times 4 = 1,008$ combinations.

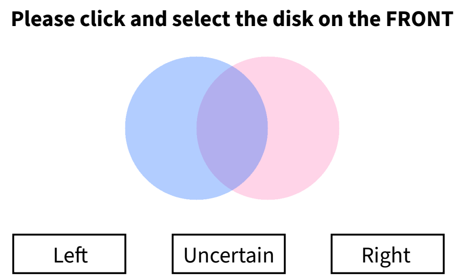

Figure 3: Task interface. Each image is 700 × 450 pixels.

Experimental design. Our pilot study tested all 1,008 combinations. Given the large number, we adopted a between-subject design to avoid fatigue effects. We randomly divided all the stimuli into four groups, each comprising 252 pairs. This allocation ensures a sufficient number of trials per participant while maintaining attention. Each participant finished one group randomly selected at the time of the experiment.

Procedure. The experiment consisted of four phases. Participants started with the standard 14 Ishihara plate test to check for color blindness [24]. They then took a training phase involving three trials. Following this, participants proceeded to the selection task, which consisted of 252 trials. We randomized these stimuli to avoid ordering effects. Finally, participants completed a demographic information questionnaire. The experiment sessions were conducted in a consistently illuminated room, with participants seated in front of a screen resolution of 2048×1080 pixels. On average, the experiment took about half an hour to complete.

Participants. We employed a convenience sampling process and recruited a total of 24 participants from our local university, including 20 males and 4 females, all of whom were Computer Science majors. All participants achieved an accuracy rate of over 90% in the color blindness test.

Main findings. We empirically observed that the error rates in depth order perception differ across the three different opacity levels. Specifically, increasing the opacity in the front layers may improve depth accuracy. In contrast, increasing opacity in the back layer leads to lower accuracy, though this effect is not statistically significant. We provide our analyses and results for the pilot study in Appx. C.

Power analysis. We utilized power analysis to determine the appropriate number of participants for our main experiment. Based on an anticipated effect size of .76 and a power level of .8, we calculated that a minimum of 15 participants would be required for each combination (each image). The effect size was derived from the pilot data, which showed an average accuracy of .79 in the task with a standard deviation of .25, compared to an expected accuracy of .6.

3.3 Main Experiment

Updates from the pilot study. We decided to further reduce the number of trials for participants, as half an hour is considered long for crowdsourced participants [2]. The stimuli were randomly divided into 9 groups, each comprising 112 trials. We also included 4 additional attention-check trials. The attention checks were presented with opaque colors, making it easy to discern the correct depth order. Furthermore, to prevent participants from spending excessive time on a single trial, we added a time limit based on the maximum average time observed, which was 20 seconds per trial. Other settings were identical to that of the pilot study: participants clicked to select the disk that appears to be in front of the others.

Participants. Following our power analysis, we recruited 192 participants via Prolific, exceeding the minimum requirement of 135 participants (9 groups × 15 participants). While there were no specific requirements for illumination, all participants were required to use a display with a resolution of at least 1920×1080 pixels. To ensure data quality, participation was restricted to U.S. residents holding at least a bachelor's degree. The demographic composition of our participant pool included 82 males, 95 females, and 3 individuals who preferred not to specify their gender. Ages ranged from 25 to 39 years, with 24 left-handed and 156 right-handed participants. Each participant was compensated $1.20 for their time as the experiment took approximately 10 minutes.

Data. We included data from participants who correctly answered at least three out of the four attention-check trials. For this reason, we excluded 12 participants, resulting in a final count of 180 participants. Each participant contributed 112 trials, adding to a total of 20,160 trials (112 trials × 180 participants). Among them, 14,768 were answered correctly (73.25%), 4,862 were answered incorrectly (24.12%), and in 530 trials, participants selected "uncertain" (2.63%).

4 Inferential Analyses and Results

We first conducted inferential analysis with the goal of drawing inferences about the tested factors. Before initiating our analyses, we formulated a set of hypotheses based on our experimental goals. Subsequently, we conducted Null Hypothesis Significance Testing (NHST) for each hypothesis. We focused on **error rates**, as completion time is not central to our interests. The **error rate** for each combination (each image) is the rate of the number of incorrect responses to the total number of responses from all participants. Because our data is not normally distributed (see Appx. D) and has variance heterogeneity, we opted for *Kruskal-Wallis H test*, a non-parametric test. We also applied *Bonferroni correction* to count for issues in multiple comparisons. Each corrected p-value is calculated by multiplying the observed p-value by the number of comparisons made, and we reported these **corrected p-values** below. To address the limitations of NHST, we also reported depth order accuracy alongside their bootstrapped confidence intervals [16, 47]. Without loss of generality, we focused our analysis on the left disks. The left and right positions are counterbalanced and symmetric (see Appx. E), and prior research confirmed no positional bias [55].

4.1 Hypotheses

H1. There are significant differences in depth order perception across various color hues, measured by error rates. Prior research suggests that cold colors in the front layer and warm colors in the back layer can enhance the depth order perception [55].[4] However, this conclusion contradicts other studies that indicate warm colors tend to appear closer [3, 53]. We hypothesize that while color hue does impact depth order perception in both the front and back layers, the effect may be less pronounced between warm and cold colors.

H2. There are significant differences in depth order perception across various opacity levels, measured by error rates. While a few studies suggest that opacity has a limited impact on depth order perception [28, 29], others argue that increasing the opacity of front objects improves the accuracy of depth order perception [1, 55]. Based on our pilot study, we anticipate that varying opacity would guide or bias participants' depth order estimates, depending on whether the opacity is applied to the front or back layers.

H3. At each opacity level, there is no significant difference in depth order perception across various color hues. Wang et al.[55] discovered that cold colors consistently resulted in better depth order perception, irrespective of opacity. We expect this finding to extend to other color hues, not just cold colors. However, we anticipate that the effect of opacity may vary depending on whether the color hue is applied to the front or back layers.

H4. In each color hue, there are significant differences in depth order perception across various opacity levels. Chan et al.[10] observed that opacity perception varied with color. We expect this finding to be applicable in our experiment as well, indicating that the effect of opacity on depth order perception may differ depending on the color hue.

4.2 The Effects of Color Hue

In Fig. 4a, we report the error rates in depth order perception across various color hues, along with confidence intervals and p-values. The Kruskal-Wallis H test indicates that the accuracy of **depth order perception is significantly dependent on color hues** (H = 20.51, p = .0046). Additionally, ● YELLOW results in significantly smaller error rates compared to others (H = 11.44, p = .0058).

We then break down the results by layer order. In either front (H = 27.08, p = .0003) or back (H = 44.26, p < .0001) layers, different color hues show significant differences. We observe that both ● PINK (H = 22.7, p < .0001) and ● YELLOW (H = 12.51, p = .0032) in *back* layers perform significantly *better*, while ● BLUE (H = 11.16, p = .0067) in *front* layers also perform significantly *better* than other color hues.

Additionally, ● GREEN, ● BLUE, and ● RED exhibit *smaller* error rates in the *front* layers compared to the *back* layers (Figs. 4b and c). In contrast, ● PURPLE, ● PINK, ● ORANGE, ● BROWN and ● YELLOW, have *smaller* error rates when they are in the *back* layers. However, only ● BLUE (H = 10.14, p = .0012) and ● PINK (H = 25.28, p < .0001) show significant differences between the front and back layers. Note that, ● RED, ● ORANGE, ● YELLOW, and ● BLUE share the same saturation value (S = 100) and nearly identical lightness values (L = {52, 51, 50, 50}), but they result in different error rates, which

[4] We use the terms "warm" and "cold" colors as a naming convention for color classification. They shouldn't be interpreted as an intention to hypothesize about color temperature.

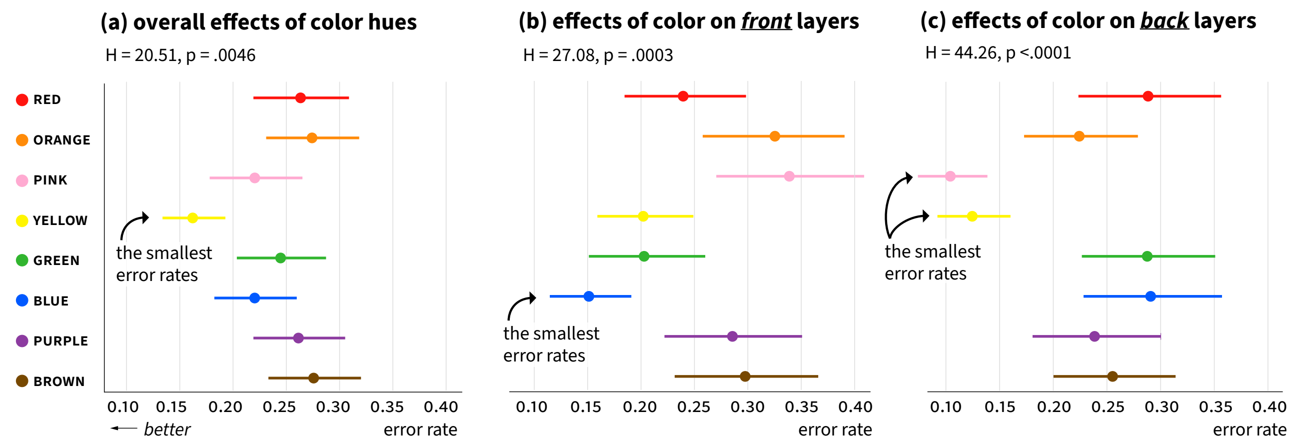

Figure 4: The effects of color hue (Sec. 4.2). We report the mean, 95% confidence intervals, and NHST results. The first panel (a) shows the overall effects of color hues aggregated across both front and back layers. After breaking down the results by (b) front and (c) back disks, we find strong effects of color hue on depth order perception.

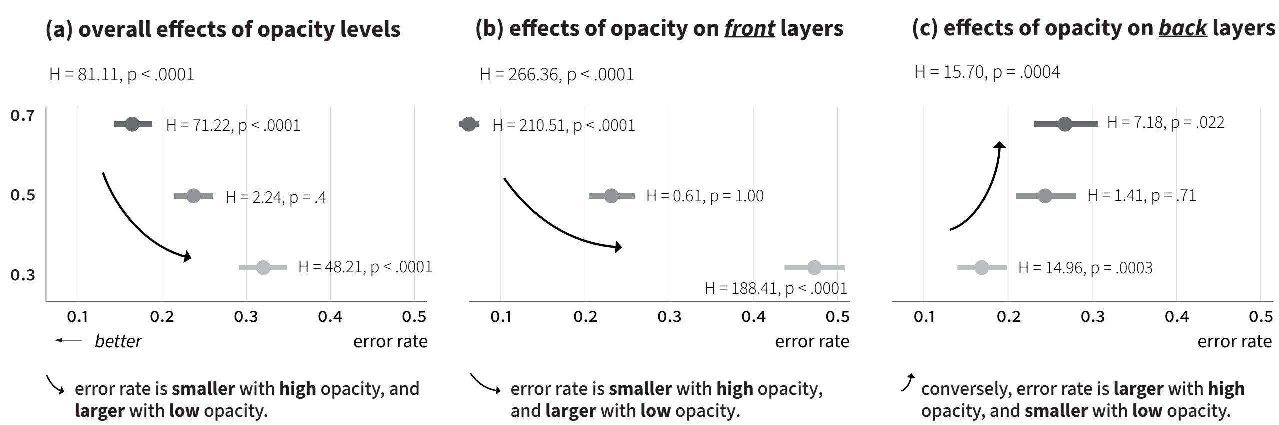

Figure 5: The effects of different opacity levels (Sec. 4.3). We report the mean, 95% confidence intervals, and NHST results. The first panel (a) shows the overall effects of opacity levels aggregated across both front and back layers. We then break down the results for (b) front disks and (c) back layers.

likely showcases the dominant effect of color hue. Readers can refer to our exploratory analysis in Sec. 6 for more details. In sum, we observe significant differences in error rates across the eight examined color hues, **fully supporting H1**.

4.3 The Effects of Opacity

In Fig. 5a, we report the error rates in depth order perception across three different opacity levels, along with the means, confidence intervals, and p-values. In general, **opacity shows significant differences** ($H = 81.11$, $p < .0001$), and we observe an overall negative correlation with the error rate. The error rate is significantly **smaller** with **high** opacity ($H = 71.22$, $p < .0001$), and **larger** with **low** opacity ($H = 48.21$, $p < .0001$).

We also break down the data by whether the disk is in the front (Fig. 5b) or back layer (Fig. 5c). When the correct disk is in the *front*, the error rate decreases as opacity increases: the error rate is significantly smaller for *high* opacity ($H = 188.41$, $p < .0001$) and smaller for high opacity ($H = 210.51$, $p < .0001$). In contrast, when a disk is in the *back*, error rates increase as opacity increases: the error rate for the back disk is significantly smaller for *low* opacity ($H = 14.96$, $p = .0003$) and *larger* for high opacity ($H = 7.18$, $p = .0022$). In sum, we observe significant differences in depth perception across various opacity levels, and depending on layer ordering, this could be a positive or negative correlation. Therefore, **H2 is fully supported**.

4.4 The Interaction Effects between Color and Opacity

a. Opacity to the Effects of Color Hue

We first categorized the effects of color hues at each opacity level (see Fig. 6a). In general, the impact of hue on depth perception diminishes as opacity increases, corresponding to smaller differences in error rates for high opacity. However, we don't observe any

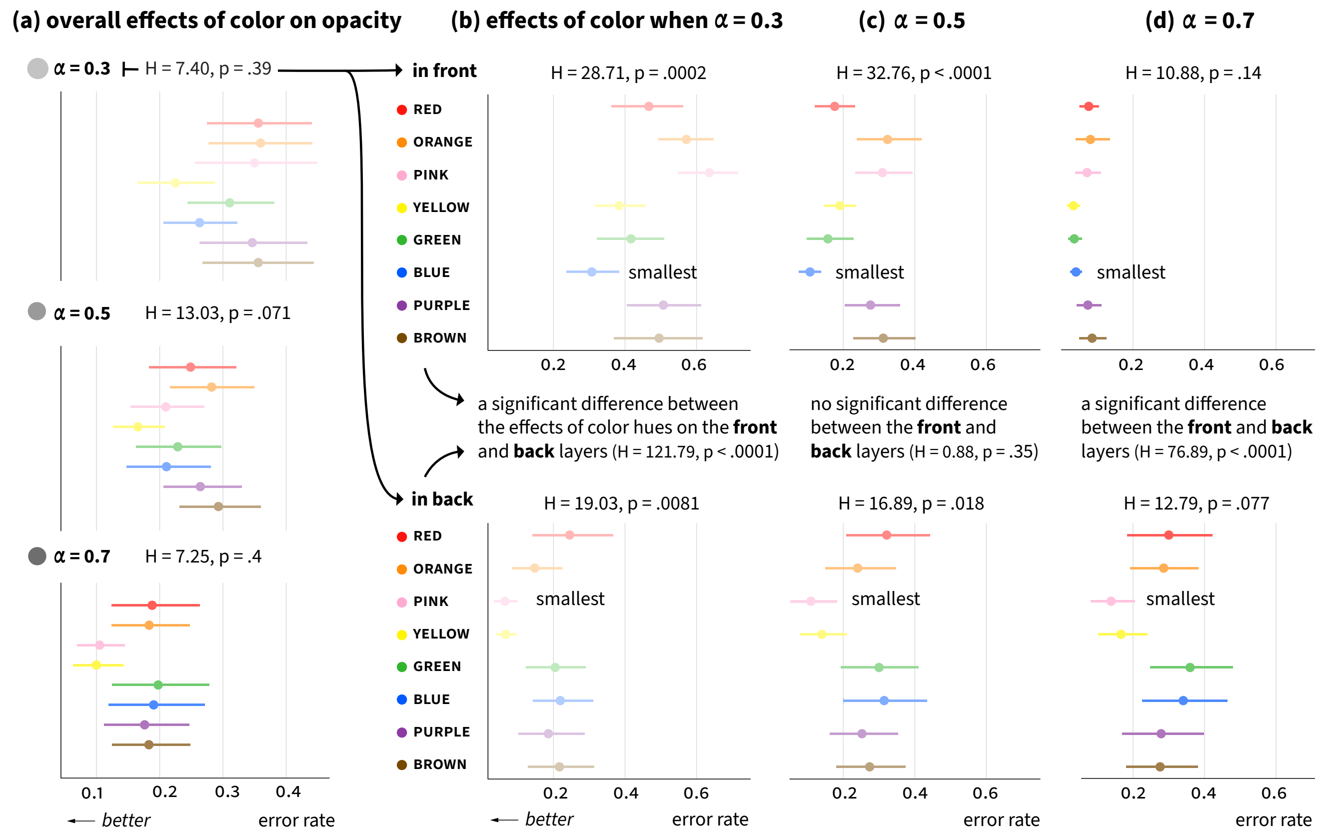

Figure 6: The interaction effects between opacity and hue (Sec. 4.4). We report mean, 95% confidence intervals, and NHST results. (a) displays the overall error rates aggregated across both front and back layers for different color hues. (b)-(d) further break down these error rates, distinguishing layer position at each of the three opacity levels.

significant differences across color hues at each examined opacity level. As before, we broke down the data by front and back layers to analyze how opacity influences the effect of colors in different layers, and delved into each opacity level as follows:

- At the low opacity level ($\alpha = 0.3$; Fig. 6b), most color hues exhibit high error rates when placed in the front layer and low error rates when placed in the back. Color hues show significantly different impacts on depth perception in both the front ($H = 28.71$, $p = .0002$) and back ($H = 19.03$, $p = .0081$) layers. The PINK color performs significantly better in the back layer ($H = 8.44$, $p = .029$) and worse in the front layer ($H = 11.29$, $p = .0062$) than other colors. Conversely, BLUE in the front layer performs significantly better than other color hues ($H = 12.26$, $p = .0037$).
- Similarly, at the medium opacity level ($\alpha = 0.5$; Fig. 6c), color hues show significantly different error rates on both the front ($H = 32.76$, $p < .0001$) and back ($H = 16.89$, $p = .0018$) layers. Again, PINK in the back layer ($H = 9.64$, $p = .0015$) and BLUE in the front layer ($H = 14.3$, $p = .0012$) demonstrate significantly smaller error rates compared to other colors.
- At the high opacity level ($\alpha = 0.7$; Fig. 6d), error rates are generally reduced for the front layer and increased for the back layer across all examined colors, and all these colors show similar error rates.

In sum, **the effect of color hues** on depth order perception **depends on opacity levels**, but this influence might be modulated by the front and back positions, **partially supporting H3**.

b. Color Hue to the Effects of Opacity

We then explored how each color hue influences the perception of depth order across the three opacity levels. We observe a significant decrease in error rates with increased opacity, particularly for PINK ($H = 12.56$, $p = .015$), YELLOW ($H = 12.55$, $p = .015$), and PURPLE ($H = 12.58$, $p = .015$). However, this trend is not significant for RED ($H = 9.5$, $p = .069$), BLUE ($H = 8.81$, $p > .098$) and GREEN ($H = 7.29$, $p > .21$).

We compared across opacity in the upper row (front layers) of Figs. 6b-d. We observed a significant reduction in error rates with an increase in opacity across all tested color hues ($p < .0001$ for all eight hues). Conversely, we compared across opacity in the lower row (back layers) of Figs. 6b-d, and observed that higher opacity might lead to higher error rates. However, this trend is not statistically significant ($p \geq .87$ for all eight hues). Overall, as the effect of opacity on depth order perception is **only** consistent among color hues in the front layer, **H4 is partially supported**.

4.5 Summary and Discussion

Overall, our results show statistically significant impacts of both color hue and opacity on depth order perception, with partial interaction effects between these two factors. Additional results detailing different pairs and layers can be found in Appxs. F and G. Our findings fully support H1 and H2, demonstrating that both color hue and opacity play a role in depth order perception. We also find that layer position is an important factor modulating the effects of color and opacity; for example, the best-performing color hues differ between the front and back layers. Compared to Wang et al.'s work [55], we did not observe the trend of all cold colors in the foreground and warm colors in the background enhancing depth order perception. This can be attributed to the complex interactions between color and opacity in semi-transparent visualizations, where simple trends may not generalize across all conditions. Our results partially support H3, suggesting that the effect of opacity on color hue's impact varies depending on whether the hue is applied to the front or back layers. Increasing front-layer opacity significantly reduces error rates across all tested color hues, while the trend of increasing back-layer opacity leading to higher error rates is not statistically significant across color hues. Unlike the cited above [55], which noted that back-layer opacity increases error rates, our findings highlight the nuanced interaction between opacity and color hue. Finally, our findings partially support H4, indicating that while color hue does influence the effect of opacity, this relationship depends on specific color hues and their layer positions. These findings largely enrich those from previous work, separately consider the impacts of hue and opacity for the front and back layers, as well as their interaction effects.

5 Predictive Analyses and Results

The inferential analyses above reveal the overall impacts of color and opacity on depth order perception. However, significance tests alone do not fully address the practical needs of designers, who may work with color choices beyond those tested in our study. To bridge this gap, we opted for machine learning models trained on our experimental data. These predictive models serve two key purposes: (1) they interpolate between tested values, enabling designers to evaluate color and opacity combinations not directly included in the experiment, which leads to our design aid tool; and (2) they allow us to analyze the importance of a set of features, providing additional insights into the nuances and relationships between design variables.

5.1 Methods

Data. Our models predict whether viewers can correctly identify the disk in the front.[5] This prediction can be used in practice to assess if viewers can correctly perceive the depth order in a visualization. As introduced in Sec. 3.3, we had collected a total of 14,768 correctly answered trials and 4,862 incorrectly answered trials from all participants. A label of 1 indicates that the participant correctly identified the front disk, while a label of 0 indicates they did not. We randomly partitioned the collected trials into training sets (90%) and test sets (10%), and used ten-fold cross-validation. The 2.63% of trials with "uncertainty" answers were discarded.

Candidate features. We derived candidate features based on the nuances of color hue and opacity between the front and back layers to effectively capture how these properties change and interact when the layers overlap. Specifically, we included the original hue and opacity values of both layers as candidate features, along with statistical measures such as minimum, maximum, and mean values to summarize the variation in hue and opacity between the two layers. Additionally, we calculated the color hue and opacity distance between the front and back layers. The hue distance refers to the difference between two colors on the hue wheel, measured as the arc-length distance [13]. In total, we derived $d = 12$ candidate features, represented as a feature vector $\boldsymbol{f} = (f^1, ..., f^d)$,

- f^1 HUE FRONT: *color hue assigned to the front layer*,
- f^2 HUE BACK: *color hue assigned to the back layer*,
- f^3 OPACITY FRONT: *opacity assigned to the front layer*,
- f^4 OPACITY BACK: *opacity assigned to the back layer*,
- f^5 HUE MIN: *the smaller hue value between the two layers*,
- f^6 HUE MAX: *the larger hue value between the two layers*,
- f^7 HUE MEAN: *the average hue value between the two layers*,
- f^8 HUE DISTANCE: *the hue distance between the two layers*,
- f^9 OPACITY MIN: *the smaller opacity between the two layers*,
- f^{10} OPACITY MAX: *the higher opacity value between the two layers*,
- f^{11} OPACITY MEAN: *the average opacity value between the two layers*, and
- f^{12} OPACITY DISTANCE: *the opacity distance between the two layers.*

Candidate architectures. In selecting model architectures, we aimed to balance the performance and interpretability, and thus included logistic regression [15], decision tree [44], random forest [7], and Support Vector Machine (SVM) with both linear and radial basis function (RBF) kernels [11, 14, 30]. First, logistic regression, despite its simplicity, has proven to be a reliable binary classification method, particularly for predicting binary judgments in scatterplots among a set of deep neural networks [59]. Decision trees offer straightforward and interpretable models by creating clear decision rules [44]. Random forests further combine the predictions of multiple decision trees with improving accuracy and robustness [7] and were the best-performing ones in making constraint-based

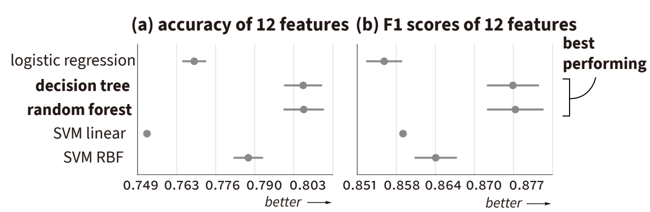

Figure 7: The average performance of different models with all 12 features. Error bars represent 95% confidence intervals. Random forest and decision tree have the best performance.

[5]We had attempted to predict error rates, but all the tested models (e.g., SVM, random forest, lasso regression) performed no better than random guessing due to noisy data or overfitting.

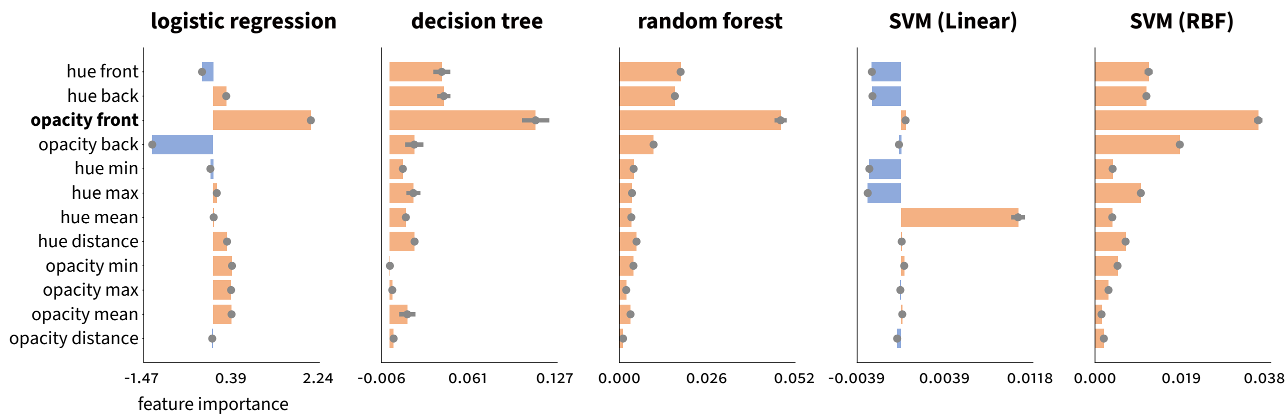

Figure 8: Feature importance extracted from the trained models. A negative value means that the feature is negatively correlated with the accuracy of depth order perception.

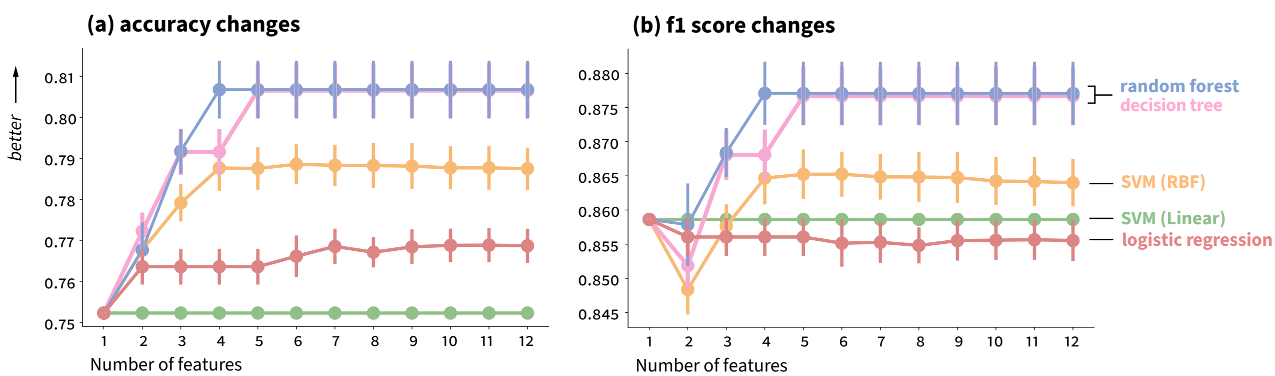

Figure 9: Sensitivity analysis. We gradually added a feature to each model by their importance and assessed the changes in performance. Again, we report mean values and 95% confidence intervals here.

visualization recommendations [30]. Finally, SVMs were used to generate optimal color blending results [31], while RBF is the most widely used kernel.

Metrics. We focused on accuracy and F1 score to assess our models. Accuracy provides an intuitive understanding of overall model performance, while the F1 score offers insights into the balance between precision and recall, especially in situations where our data exhibits class imbalances. We also calculated other common metrics like precision and recall and provided the results in Appx. H.

5.2 Results

Performance. We report the mean accuracy and F1 scores, along with their 95% confidence intervals in Figs. 7a-b. Overall, the decision tree and random forest models achieve the best performance among those tested. Between them, random forest models result in slightly higher prediction accuracy and F1 score compared to decision tree. On the other hand, SVM with a linear kernel and logistic regression perform relatively poorly, but are still close to SVM with an RBF kernel.

Feature importance. We also calculated feature importance for each model, and reported the results in Fig. 8. For logistic regression and SVM with a linear kernel, the coefficients were directly extracted from the linear models. For the nonlinear models of decision tree, random forest, and SVM with an RBF kernel, we employed the permutation importance scheme [7] to assess feature importance. This method measures the model's performance changes when feature values are randomly reshuffled. Across all tested models, except SVM with a linear kernel, the top feature is OPACITY FRONT, the opacity of the front layer. In logistic regression, all of the top five most important features are related to opacity. The top features include those related to hue in other models.

Sensitivity analysis. We also conducted a sensitivity analysis. Specifically, we selected the top n ($n \leq 12$) candidate features and retrained the models. We reported the accuracy and F1 score, along with 95% confidence intervals, as a function of the number of top features in Fig. 9. While most of them showed increasing accuracy as more features were added, the SVM with linear kernels exhibited no performance improvement beyond the top feature, HUE MEAN. Additionally, the decision tree and random forest models

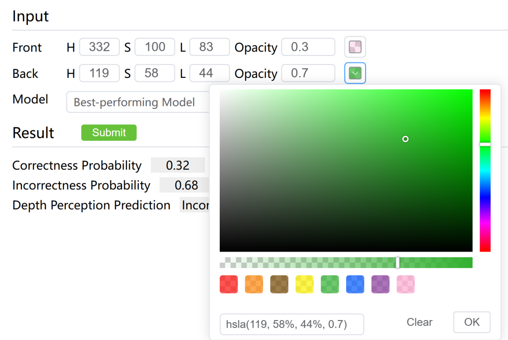

Figure 10: Interface of our small design aid tool. Appx. I provides a demonstration.

achieved the highest accuracy with the top features, and features like HUE MAX *(the larger hue between two layers)*, OPACITY MAX *(the larger opacity between two layers)*, and OPACITY DISTANCE *(the opacity distance between two layers)* were redundant.

Best-performing model. The random forest with four features (OPACITY FRONT, HUE FRONT, HUE BACK, OPACITY BACK) is considered the best-performing model, as it consistently yielded the highest average accuracy and F1 score. These features suggest a balance of color hue and opacity in both the front and back layers.

Design Tool. Since these predictive models yield satisfying accuracy, we applied them in a small tool that can aid in the design of semi-transparent visualizations. In this small tool, a user can input the color, opacity, and layer positions of two overlapping elements in the interface, and it will **automatically output whether viewers can correctly perceive the depth order** with a probability score (see Fig. 10). The design tool, our code, and models are available at https://osf.io/n3jg8, including detailed instructions for downloading and running the tool on a local computer.

6 Exploratory Analyses

As reported in previous work [26, 55], lightness and saturation may influence depth perception. As a robustness check, we **replicated** our inferential and predictive analyses for the lightness and saturation of the eight colors used in the experiment. As these analyses were conducted post hoc, we consider them exploratory. We summarize the findings below and report details in Appxs. J and K.

Lightness. The Kruskal-Wallis H test on lightness results in a significant difference ($H = 17.26, p = 0.0084$) for both front ($H = 25.23$, $p = .0003$) and back ($H = 28.90$, $p < .0001$) layers. As lightness varies hue in the tested colors, we then used predictive analysis to separate the effect of lightness from opacity and hue. We used the original candidate feature vector f augmented with six additional lightness features, and calculated feature importance across the five models. Across all models, opacity- and hue-related features consistently ranked highest. Hue-related features dominate the top five ranks, except in the logistic regression models. Also, adding lightness features minimally improved model accuracy (on average less than 0.16%). In sum, lightness has minimal influence compared to opacity and hue, validating our results and conclusions in Secs. 4 and 5.

Saturation. The Kruskal-Wallis H tests show no significant differences in depth order perception across different saturation values ($H = 1.39, p = 0.5$), with consistent results observed for both the front ($H = 5.22, p = 0.074$) and back ($H = 5.08, p = 0.079$) layers. Again, opacity- and hue-related features consistently ranked highest, and no saturation features appeared in the top five across all the tested models. Furthermore, adding saturation features also minimally improved model accuracy (on average about 0.2%). In sum, saturation also has minimal influence.

7 General Discussion

7.1 Design Guidelines

The selection of colors and opacity, as well as their combinations, are composed of a large design space. A casual design might result in incorrect depth order perception. Based on our inferential and predictive analyses, we provide four design guidelines for color selection in semi-transparent visualizations as follows:

- Color hue significantly affects the perception of depth layers. Among the most commonly named colors, we found that PINK and YELLOW perform superior in depth order perception compared to other colors in the back layer. However, in the front layer BLUE performs better than other colors. Furthermore, increasing the hue distance between the front and back layers has a positive effect on depth order perception, according to the best-performing model. Therefore, GL1 **BLUE and PINK (or YELLOW) can be considered a priority in selecting colors for depth order perception tasks in the *front* and *back* layers, respectively**.
- More broadly, warm hues (e.g.,) and cold hues (e.g.,) have different impacts on depth order perception. Specifically, disks with GREEN, BLUE, and RED have *smaller* error rates when positioned in the *front* layer as opposed to the *back* layer. Conversely, when disks containing PURPLE, PINK, ORANGE, BROWN, and YELLOW are in the *back* layer as opposed to the front layer, they exhibit *smaller* error rates. Previous suggestions by Wang et al. [55] and others have posed that cold colors should be in front and warm colors should be in the back. Although RED and PURPLE present exceptions to this pattern, in most cases, our findings support the idea that warm and cold colors can be used as a reference for depth order perception. As such, GL2 **cold colors can be used for the front layer and warm colors for the back layer**, except RED and PURPLE.
- Opacity significantly influences depth perception, and the impact of opacity on depth order perception remains consistent across varying color hues. However, the effect of opacity is conditional on the layer positions, whether in the back or the front layer. For disks in the back layer, higher opacity levels are associated with increased errors in depth perception. In contrast, for disks in the front layer, lower opacity levels can lead to greater errors. The importance of opacity in both the front and back layers is also evident in our predictive analyses, with

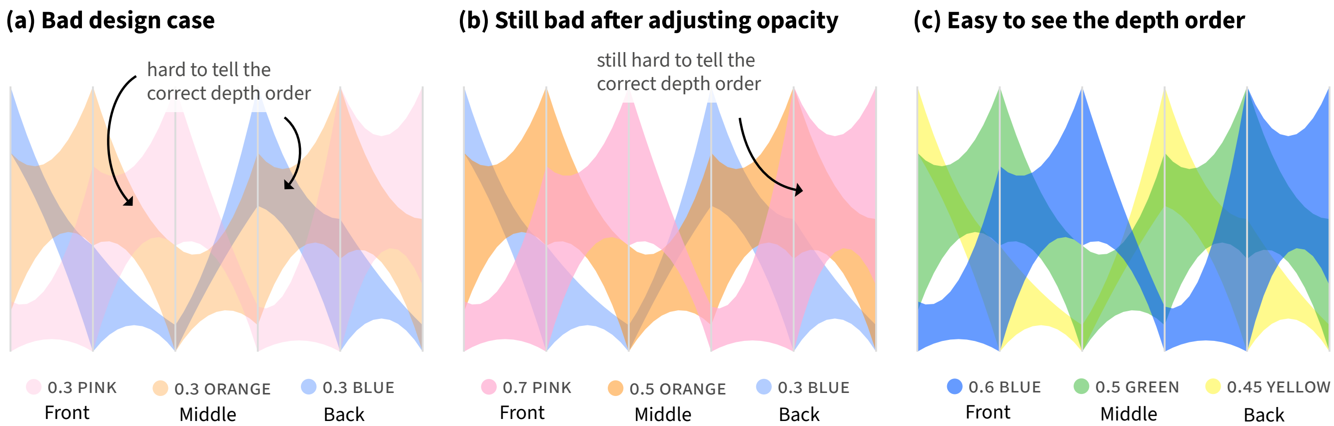

Figure 11: Applying our design guidelines to the Scan Bio dataset [49]. (a) and (b) violate GL1 to put ● PINK in front layers and ● BLUE in back layers, resulting in ambiguity in order perception regardless of opacity. When we follow GL1 and GL2 to put ● BLUE and cold colors in front layers, and use GL3 to adjust opacity, the depth order is easy to perceive in (c).

the opacity of the front layer showing a stronger effect. As such, we recommend that GL3 **front layers should be assigned to higher opacity, while back layers should be assigned to lower opacity**. However, as opacity levels further increase or decrease, both depth perception and overlapping area identification tasks become more difficult outside our tested range (e.g., opacity > 0.7). Careful calibration of these opacity levels is essential to achieve a balance that optimizes both occlusion effects and depth order perception.

- The effect of opacity on depth perception varies with color hues, in alignment with the suggestions of [10]. As opacity increases, the error rates decrease for colors like ● PINK, ● YELLOW, and ● PURPLE. But this trend is not significant for other colors like ● RED, ● BLUE and ● GREEN. Therefore, GL4 **in applications that require frequently adjusting opacity levels—such as medical imaging for exploring different organ structures—using colors like ● RED, ● BLUE and ● GREEN can be more effective.** These colors (●●●) tend to produce fewer errors compared to others when opacity is frequently changed.

Additionally, the influence of color hue on depth perception has a diminishing trend as opacity increases, but the effect is not significant. Although we observed that the cold color ● BLUE performs better in the front layer than other colors in both low and medium opacity levels, we cannot derive that cold colors consistently result in better depth order perception regardless of opacity as Wang et al. [55] suggested, due to the interaction effects between color and opacity.

We demonstrate the practical application of our design guidelines using a case of parallel coordinates with three overlapping clusters. We show two negative examples in Figs. 11a and b, where the front clusters are assigned ● PINK to front layers and ● BLUE to back layers. Regardless of whether the opacity level is higher or lower, it is difficult to perceive the correct depth order. Let's first follow GL1 and assign ● BLUE to the front layer; we then follow GL3 and decrease the opacity from the front to the back. As a result, it is much easier to perceive the correct depth order and identify the clusters in Fig. 11c. These examples demonstrate the generalizability of our guidelines to situations involving various shapes and more complex color overlaps, even though our experiment only involves two overlapping disks.

7.2 Limitations & Future Work

Our experiment represents the most comprehensive effort to date in evaluating the impacts of color hue and opacity on depth order perception, yet we had to simplify our experiment to a manageable scope.

First, our study opted to use α-blending, to ensure the relevance and applicability of our findings to a broader audience. Future studies could expand on our findings by alternative methods [12, 55] to address specific depth perception challenges. Due to color blending, background colors may alter the appearance of foreground colors, leading to different depth layer perceptions (see Appx. L). We also hope our results will inspire the development of color blending operators [12, 31] that better account for the perceptual interactions between color and opacity.

Second, other visual elements, such as color saturation [10], color lightness [19], shape [32], size [50], x-junctions, boundaries, and texture [34, 42], could also influence depth perception. Our study provides a foundation for future studies to incorporate these additional factors. Similarly, we opted for simpler and interpretable predictive models based on prior works, while future work could explore the nuances in accuracy and F1 score for different architectures.

Third, our experiment used a controlled setup with two overlapping disks and generated interpretable results that provide useful insights for future work. More complex, multi-layered scenarios could be included in future studies to improve the realistic visual experiences and investigate the wider applicability of our discoveries. On the other hand, our chosen colors primarily reflect how humans categorize colors in practical visualization applications, but were not controlled for lightness and saturation. Although our exploratory analyses show minimal impacts of lightness and saturation, follow-up studies with controlled lightness and saturation could support fine-grained adjustments of color hue and opacity values.

Finally, we mainly focus on understanding the effect of color and opacity on depth perception, a foundation for many tasks in semi-transparent visualizations. There are other meaningful tasks in semi-transparent visualization, such as identifying overlapping areas across different categories and structures. Further research could explore these different tasks to expand the applicability of our findings to a broader range of domain applications.

8 Conclusion

Our goal was to investigate the impact of color hue and opacity on depth perception in semi-transparent visualizations. To achieve this, we conducted an online experiment, testing semi-transparent visualizations consisting of two disks that varied across eight color hues, three opacity levels, two layer orders, and two arrangements. We analyzed the experimental data using inferential analysis, which highlighted the significant influence of opacity and other factors on depth perception. Additionally, we performed predictive analyses with 12 candidate features to evaluate their impact on depth order and develop predictive depth models that can be easily applied in practical scenarios. In this analysis, both decision tree and random forest achieved comparable results, with the best average accuracy of 80.72% and F1 score of 87.75% on the tested color hue and opacity features. Based on both inferential and predictive analyses, we provided a small design tool and derived four design guidelines for semi-transparent visualizations. We hope our work can inspire and encourage further research in the field of semi-transparent visualizations.

Acknowledgments

We would like to thank Weifeng Ni for his contributions to data collection. In this study, Qiong Zeng is supported by grants from the National Key Research and Development Program of China (No. 2021YFF0704300) and the National Natural Science Foundation of China (No. 62372271); Lin Lu is supported by the National Natural Science Foundation of China (No. 62472258); and Yunhai Wang is supported by the National Natural Science Foundation of China (No. 62132017), the Natural Science Foundation of Shandong Province (No. ZQ2022JQ32), the Fundamental Research Funds for the Central Universities, the Natural Science Foundation of Beijing Municipality (No. L247027), and the Research Funds of Renmin University of China.

References

[1] Edward H. Adelson and P. Anandan. 1990. Ordinal characteristics of transparency. In *AAAI Workshop on Qualitative Vision*. Boston, MA, USA.

[2] Daniel Archambault, Helen C Purchase, and Tobias Hoßfeld. 2017. *Evaluation in the Crowd*. Springer. https://doi.org/10.1007/978-3-319-66435-4

[3] Reynold Bailey, Cindy Grimm, and Christopher Davoli. 2006. The effect of warm and cool object colors on depth ordering. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*. Association for Computing Machinery, New York, NY, USA, 161. https://doi.org/10.1145/1140491.1140538

[4] Lyn Bartram, Billy Cheung, and Maureen Stone. 2011. The Effect of Colour and Transparency on the Perception of Overlaid Grids. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 1942–1948. https://doi.org/10.1109/TVCG.2011.242

[5] Brent Berlin and Paul Kay. 1969. *Basic Color Terms: Their Universality and Evolution*. University of California Press. https://doi.org/10.2307/412128

[6] Christian Boucheny, Georges-Pierre Bonneau, Jacques Droulez, Guillaume Thibault, and Stephane Ploix. 2009. A perceptive evaluation of volume rendering techniques. *ACM Transactions on Applied Perception* 5, 4 (2009). https://doi.org/10.1145/1462048.1462054

[7] Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (2001), 5–32. https://doi.org/10.1023/A:1010933404324

[8] Roxana Bujack, Terece L. Turton, Francesca Samsel, Colin Ware, David H. Rogers, and James Ahrens. 2018. The Good, the Bad, and the Ugly: A Theoretical Framework for the Assessment of Continuous Colormaps. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 923–933. https://doi.org/10.1109/TVCG.2017.2743978

[9] Michael Burns, Martin Haidacher, Wolfgang Wein, Ivan Viola, and M. Eduard Gröller. 2007. Feature emphasis and contextual cutaways for multimodal medical visualization. In *Proceedings of the 9th Joint Eurographics / IEEE VGTC Conference on Visualization (EUROVIS'07)*. Eurographics Association, Goslar, DEU, 275–282. https://doi.org/10.2312/VisSym/EuroVis07/275-282

[10] Ming-Yuen Chan, Yingcai Wu, Wai-Ho Mak, Wei Chen, and Huamin Qu. 2009. Perception-Based Transparency Optimization for Direct Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1283–1290. https://doi.org/10.1109/TVCG.2009.172

[11] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011). https://doi.org/10.1145/1961189.1961199

[12] Johnson Chuang, Daniel Weiskopf, and Torsten Moller. 2009. Hue-Preserving Color Blending. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 1275–1282. https://doi.org/10.1109/TVCG.2009.150

[13] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. 2006. Color harmonization. *ACM Transactions on Graphics* 25, 3 (2006), 624–630. https://doi.org/10.1145/1141911.1141933

[14] Corinna Cortes and Vladimir Vapnik. 1995. Support-Vector Networks. *Machine Learning* 20, 3 (1995), 273–297. https://doi.org/10.1023/A:1022627411411

[15] David Roxbee Cox. 1958. The Regression Analysis of Binary Sequences. *Journal of the Royal Statistical Society: Series B (Methodological)* 20, 2 (1958), 215–232. https://doi.org/10.1111/j.2517-6161.1958.tb00292.x

[16] Pierre Dragicevic. 2016. Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*. Springer, 291 – 330. https://doi.org/10.1007/978-3-319-26633-6_13

[17] Rickard Englund and Timo Ropinski. 2016. Evaluating the perception of semi-transparent structures in direct volume rendering techniques. In *SIGGRAPH ASIA 2016 Symposium on Visualization (SA '16)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3002151.3002164

[18] Rickard Englund and Timo Ropinski. 2018. Quantitative and Qualitative Analysis of the Perception of Semi-Transparent Structures in Direct Volume Rendering. *Computer Graphics Forum* 37, 6 (2018), 174–187. https://doi.org/10.1111/cgf.13320

[19] Shi-Min Gong, Fan-Yu Liou, and Wen-Yuan Lee. 2023. The effect of lightness, chroma, and hue on depth perception. *Color Research & Application* 48, 6 (2023), 793–800. https://doi.org/10.1002/col.22894

[20] Alba Grieco and Sergio Roncato. 2005. Lines That Induce Phenomenal Transparency. *Perception* 34, 4 (2005), 391–407. https://doi.org/10.1068/p5142

[21] Haleh Hagh-Shenas, Sunghee Kim, Victoria Interrante, and Christopher Healey. 2007. Weaving Versus Blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1270–1277. https://doi.org/10.1109/TVCG.2007.70623

[22] Lane Harrison, Fumeng Yang, Steven Franconeri, and Remco Chang. 2014. Ranking Visualizations of Correlation Using Weber's Law. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 1943–1952. https://doi.org/10.1109/TVCG.2014.2346979

[23] Jeffrey Heer and Maureen Stone. 2012. Color Naming Models for Color Selection, Image Editing and Palette Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1007–1016. https://doi.org/10.1145/2207676.2208547

[24] Shinobu Ishihara. 1951. *Tests for Colour-blindness*. Nippon Isho Shuppan Company. https://doi.org/10.1093/oxfordjournals.bmb.a074317

[25] James Jaccard, Robert Turrisi, and Choi K Wan. 1990. *Interaction effects in multiple regression*. Sage Publications. Includes bibliographical references (pages 90-94).

[26] Jimmy Johansson, Patric Ljung, and Matthew Cooper. 2007. Depth Cues and Density in Temporal Parallel Coordinates. In *Eurographics/ IEEE-VGTC Symposium on Visualization*, K. Museth, T. Moeller, and A. Ynnerman (Eds.). The Eurographics Association. https://doi.org/10.2312/VisSym/EuroVis07/035-042

[27] Daniel Jönsson, Erik Sundén, Anders Ynnerman, and Timo Ropinski. 2014. A Survey of Volumetric Illumination Techniques for Interactive Volume Rendering. *Computer Graphics Forum* 33, 1 (2014), 27–51. https://doi.org/10.1111/cgf.12252

[28] Daniel Kersten, Heinrich H. Bülthoff, Bennett L. Schwartz, and Kenneth J. Kurtz. 1992. Interaction between transparency and structure from motion. *Neural Computation* 4, 4 (1992), 573–589. https://doi.org/10.1162/neco.1992.4.4.573

[29] Marta Kersten, James Stewart, Niko Troje, and Randy Ellis. 2006. Enhancing Depth Perception in Translucent Volumes. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 1117–1124. https://doi.org/10.1109/TVCG.2006.139

[30] Hyeok Kim, Ryan Rossi, Abhraneel Sarma, Dominik Moritz, and Jessica Hullman. 2022. An Automated Approach to Reasoning About Task-Oriented Insights in Responsive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 129–139. https://doi.org/10.1109/TVCG.2021.3114782

[31] Lars Kühne, Joachim Giesen, Zhiyuan Zhang, Sungsoo Ha, and Klaus Mueller. 2012. A Data-Driven Approach to Hue-Preserving Color-Blending. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2122–2129. https://doi.org/10.1109/TVCG.2012.186

[32] Éric Languénou. 2023. Contrast Driven Color-Group Assignment in Categorical Data Visualization. In *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications VISIGRAPP*, Christophe Hurter, Helen C. Purchase, and Kadi Bouatouch (Eds.). SCITEPRESS, 53–64. https://doi.org/10.5220/0011615000003417

[33] Kecheng Lu, Mi Feng, Xin Chen, Michael Sedlmair, Oliver Deussen, Dani Lischinski, Zhanglin Cheng, and Yunhai Wang. 2021. Palettailor: Discriminable Colorization for Categorical Data. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 475–484. https://doi.org/10.1109/TVCG.2020.3030406

[34] Martin Luboschik, Axel Radloff, and Heidrun Schumann. 2010. A new weaving technique for handling overlapping regions. In *Proceedings of the International Conference on Advanced Visual Interfaces* (Roma, Italy) *(AVI '10)*. Association for Computing Machinery, New York, NY, USA, 25–32. https://doi.org/10.1145/1842993.1842999

[35] Deng Luo, Zainab Alsuwaykit, Dawar Khan, Ondřej Strnad, Tobias Isenberg, and Ivan Viola. 2025. DiffFit: Visually-Guided Differentiable Fitting of Molecule Structures to a Cryo-EM Map. *IEEE Transactions on Visualization and Computer Graphics* 31 (2025). https://doi.org/10.48550/arXiv.2404.02465

[36] Justin Matejka, Fraser Anderson, and George Fitzmaurice. 2015. Dynamic Opacity Optimization for Scatter Plots. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2707–2710. https://doi.org/10.1145/2702123.2702585

[37] Adrian Mayorga and Michael Gleicher. 2013. Splatterplots: Overcoming Overdraw in Scatter Plots. *IEEE Transactions on Visualization and Computer Graphics* 19, 9 (2013), 1526–1538. https://doi.org/10.1109/TVCG.2013.65

[38] Caitlyn M. McColeman, Fumeng Yang, Timothy F. Brady, and Steven Franconeri. 2022. Rethinking the Ranks of Visual Channels. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 707–717. https://doi.org/10.1109/TVCG.2021.3114684

[39] Luana Micallef, Gregorio Palmas, Antti Oulasvirta, and Tino Weinkauf. 2017. Towards Perceptual Optimization of the Visual Design of Scatterplots. *IEEE Transactions on Visualization and Computer Graphics* 23, 6 (2017), 1588–1599. https://doi.org/10.1109/TVCG.2017.2674978

[40] Tamara Munzner. 2014. *Visualization Analysis and Design*. CRC Press. https://doi.org/10.1201/b17511

[41] Brian D. Ondov, Fumeng Yang, Matthew Kay, Niklas Elmqvist, and Steven Franconeri. 2021. Revealing Perceptual Proxies with Adversarial Examples. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1073–1083. https://doi.org/10.1109/TVCG.2020.3030429

[42] Arpha Pisanpeeti and Éric Dinet. 2017. Transparent objects: Influence of shape and color on depth perception. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1867–1871. https://doi.org/10.1109/ICASSP.2017.7952480

[43] Thomas Porter and Tom Duff. 1984. Compositing digital images. *SIGGRAPH Computer Graphics* 18, 3 (1984), 253–259. https://doi.org/10.1145/964965.808606

[44] John Ross Quinlan. 1985. Introduction of Decision Trees. *Machine Learning* 1 (1985), 81–106. https://doi.org/10.1007/BF00116251

[45] Khairi Reda, Amey A. Salvi, Jack Gray, and Michael E. Papka. 2021. Color Nameability Predicts Inference Accuracy in Spatial Visualizations. *Computer Graphics Forum* 40, 3 (2021), 49–60. https://doi.org/10.1111/cgf.14288

[46] Khairi Reda and Danielle Albers Szafir. 2020. Rainbow Revisited: Modeling Effective Colormap Design for Graphical Inference. *IEEE Transactions on Visualization and Computer Graphics* (2020). https://doi.org/10.1109/TVCG.2020.3030439

[47] Judy Robertson and Maurits Kaptein. 2018. *Modern Statistical Methods for HCI*. Springer Publishing Company. https://doi.org/10.1007/978-3-319-26633-6

[48] Yuichi Sakano, Yurina Kitaura, Kyoko Hasegawa, Roberto Lopez-Gulliver, Liang Li, Hiroshi Ando, and Satoshi Tanaka. 2018. Quantitative evaluation of perceived depth of transparently-visualized medical 3D data presented with a multi-view 3D display. *International Journal of Modeling, Simulation, and Scientific Computing* 09, 03 (2018), 1840009. https://doi.org/10.1142/S1793962318400093

[49] Christopher A. Shoemaker2001, Manish Pungliya, Michael A. Sao Pedro, Carolina Ruiz, Sergio A. Alvarez, Matthew Ward, Elizabeth F. Ryder, and Julia Krushkal. 2001. Computational Methods for Single-Pointand Multipoint Analysis of GeneticVariants Associated with a SimulatedComplex Disorder in a GeneralPopulation. *Genetic Epidemiology* 21, 1 (2001), 738–745. https://doi.org/10.1002/gepi.2001.21.s1.s738

[50] Maureen Stone. 2012. In Color Perception, Size Matters. *IEEE Computer Graphics and Application* 32, 2 (2012), 8–13. https://doi.org/10.1109/MCG.2012.37

[51] Danielle Albers Szafir. 2018. Modeling Color Difference for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 392–401. https://doi.org/10.1109/TVCG.2017.2744359

[52] Christian van Onzenoodt, Anke Huckauf, and Timo Ropinski. 1978. Effects of colour on perceived depth. Review of experiments and evaluation of theories. *Scandinavian Journal of Psychology* 19 (1978), 133—-143. Issue 2. https://doi.org/10.1111/j.1467-9450.1978.tb00313.x

[53] Christian van Onzenoodt, Anke Huckauf, and Timo Ropinski. 2020. On the Perceptual Influence of Shape Overlap on Data-Comparison Using Scatterplots. *Computers & Graphics* 90 (2020), 169–181. Issue 5. https://doi.org/10.1016/j.cag.2020.05.028

[54] Chaoli Wang, Hongfeng Yu, and Kwan-Liu Ma. 2008. Importance-Driven Time-Varying Data Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1547–1554. https://doi.org/10.1109/TVCG.2008.140

[55] Lujin Wang, Joachim Giesen, Kevin T. McDonnell, Peter Zolliker, and Klaus Mueller. 2008. Color Design for Illustrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1739–1754. https://doi.org/10.1109/TVCG.2008.118

[56] Colin Ware. 2004. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, San Francisco, CA, USA. https://doi.org/10.1016/C2009-0-62432-6

[57] Edward J. Wegman and Qiang Luo. 1996. High Dimensional Clustering Using Parallel Coordinates and the Grand Tour. *Computing Science and Statistics* 28, 1 (1996), 352–360. https://doi.org/10.1007/978-3-642-59051-1_10

[58] Chaoqing Xu, Guodao Sun, and Ronghua Liang. 2021. A survey of volume visualization techniques for feature enhancement. *Visual Informatics* 5, 3 (2021), 70–81. https://doi.org/10.1016/j.visinf.2021.08.001

[59] Fumeng Yang, Yuxin Ma, Lane Harrison, James Tompkin, and David H. Laidlaw. 2023. How Can Deep Neural Networks Aid Visualization Perception Research? Three Studies on Correlation Judgments in Scatterplots. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 822. https://doi.org/10.1145/3544548.3581111

[60] Qiong Zeng, Yongwei Zhao, Yinqiao Wang, Jian Zhang, Yi Cao, Changhe Tu, Ivan Viola, and Yunhai Wang. 2022. Data-Driven Colormap Adjustment for Exploring Spatial Variations in Scalar Fields. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2022), 4902–4917. https://doi.org/10.1109/TVCG.2021.3109014

[61] Lin Zheng, Yingcai Wu, and Kwan-Liu Ma. 2013. Perceptually-Based Depth-Ordering Enhancement for Direct Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics* 19, 3 (2013), 446–459. https://doi.org/10.1109/TVCG.2012.144

[62] Liang Zhou and Charles D. Hansen. 2016. A Survey of Colormaps in Visualization. *IEEE Transactions on Visualization and Computer Graphics* 22, 8 (2016), 2051–2069. https://doi.org/10.1109/TVCG.2015.2489649