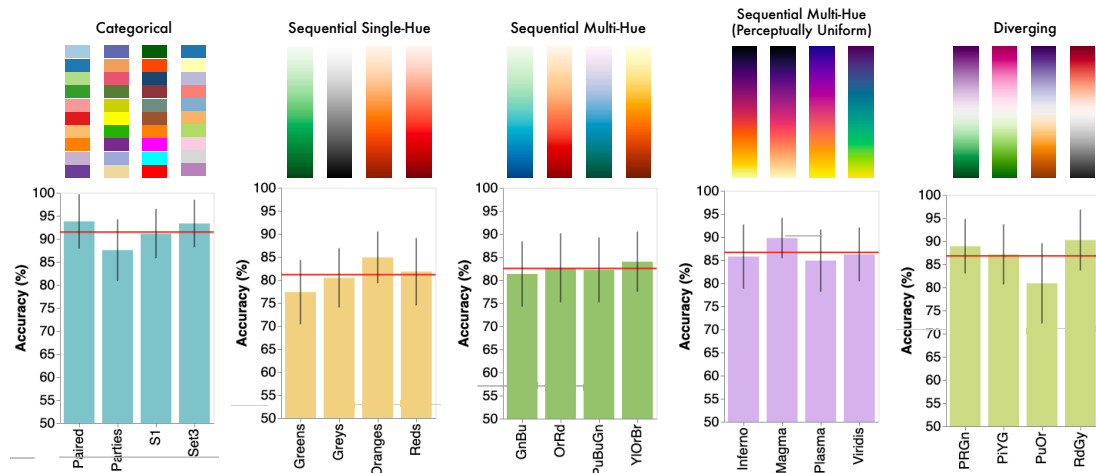# Revisiting Categorical Color Perception in Scatterplots: Sequential, Diverging, and Categorical Palettes

Chin Tseng[1], Arran Zeyu Wang[1], Ghulam Jilani Quadri[1,2], and Danielle Albers Szafir[1]

[1]University of North Carolina at Chapel Hill, NC, USA
[2]University of Oklahoma, OK, USA

Figure 1: *The accuracy rates for relative mean judgment tasks separated per color palette design family, with their average accuracy per family (red line). We find that design families significantly influenced class comparison, with multi-hue categorical palettes providing the highest performance (91.44%), followed by diverging (86.78%), perceptually-uniform sequential palettes (86.67%), multi-hue sequential palettes (82.56%), and single-hue sequential palettes (81.11%).*

Abstract

Existing guidelines for categorical color selection are heuristic, often grounded in intuition rather than empirical studies of readers' abilities. While design conventions recommend palettes maximize hue differences, more recent exploratory findings indicate other factors, such as lightness, may play a role in effective categorical palette design. We conducted a crowdsourced experiment on mean value judgments in multi-class scatterplots using five color palette families—single-hue sequential, multi-hue sequential, perceptually-uniform multi-hue sequential, diverging, and multi-hue categorical—that differ in how they manipulate hue and lightness. Participants estimated relative mean positions in scatterplots containing 2 to 10 categories using 20 colormaps. Our results confirm heuristic guidance that hue-based categorical palettes are most effective. However, they also provide additional evidence that scalable categorical encoding relies on more than hue variance.

CCS Concepts

• Human-centered computing → Information visualization; Empirical studies in visualization;

1. Introduction

People use scatterplots for a range of tasks [SHGF16, QR21, Qua21], including comparing trends [HYFC14], exploring clusters [QR20, QNWR22, JQL*23], and estimating values [GCNF13, HWS21]. Designers commonly use mark color to encode categorical data in scatterplots [GCNF13]. However, categories become harder to compare as the number of categories increases [TQWS23]. Intentional design choices can provide robust data interpretation as the number of categories increases; however, most palette design guidelines are heuristic, often grounded in designer intuition rather than quantitative evidence.

One common guideline for categorical palette design is to maximize hue variation [Mun14, GLS16, RNAK18, Sto06]. Tools like ColorBrewer [HB03] or Tableau offer categorical encodings that embody this idea, first emphasizing nameable color differences, then varying lightness and saturation as secondary factors to extend palettes to larger numbers of colors. The principles of effectiveness and expressiveness [Mun14, Mac86] further suggest that palettes emphasize hue variation over lightness to avoid unintentionally implying order within a palette and to align with human intuition. However, recent studies indicate that these principles may not apply universally or uniformly [TQWS23]. Instead, they con-

tend lightness variation can strengthen categorical perception, but draw these conclusions by studying only palettes with strong baseline hue variation. This study aims to examine traditional design heuristics for categorical color encoding empirically.

We investigate these heuristics in multiclass scatterplots, measuring performance across five design families of palettes characterized by their respective use of hue and lightness: categorical, single-hue sequential, multi-hue sequential, perceptually-uniform multi-hue sequential, and diverging (Figure 1). Classical visualization guidelines encourage the use of hue-varying categorical palettes and discourage other families that rely more heavily on lightness. However, recent exploratory analyses found that hue variation alone may not be the strongest predictor of robustness in categorical color encoding design [TQWS23]. Our approach explores this tension in a class averaging task previously used to investigate the role of color. We employed the task of judging the highest mean y-values in multiclass scatterplots [GCNF13, TQWS23].

We conducted a crowdsourced experiment to explore the impact of sequential, diverging, and categorical color palettes on categorical perception. Our experimental results reveal that both enlarging hue variance (two-hue diverging or multi-hue sequential) and emphasizing perceptual uniformity can benefit categorical perception and lead to more robust performance over larger numbers of categories. We reconcile our results with current empirical design guidelines grounded in studies of exclusively categorical palettes [TQWS23], concluding that while hue variation dominates performance, lightness variation and general perceptual distance also play a notable role in categorical encoding design.

2. Related Work

We briefly review related literature about categorical perception in scatterplots and color palette design and perception.

2.1. Categorical Perception in Scatterplots

Categorical perception [GH10] describes how people perceive different groups (or categories) of objects. Existing studies assessed categorical perception in scatterplots for various tasks, such as measuring the influence of shape, size, and color [Sza18, SS19, BSG17], and approximating difference metrics between categorical channels [DBH14]. However, the emphasis on minimum differentiable thresholds in these studies characterizes pairwise interactions between marks; the results may not generalize to holistic relationships across an entire palette.

Subsequent studies have explored people's abilities to compare values between larger numbers of categories. An initial experiment focused on comparing mean values in scatterplots, indicating that performance remained consistent as the number of categories increased from two to three [GCNF13]. More recently, Tseng et al. [TQWS23] further extended the mean estimation task to more complex scatterplots with up to ten categories, finding that both the number of categories and discriminability of colors significantly impact people's abilities to compare categories. However, these studies focused on a limited number of colors with an emphasis on color palettes that primarily varied in hue. Tseng et al.'s exploratory analysis suggested that lightness may play a larger role than hue in predicting palette robustness across varying category numbers. We

extend this observation to consider a wider range of color design choices that vary along both hue and lightness.

2.2. Color Palette Design and Perception

Prior work offers a range of design guidelines and measures for effective colormap design [BTS*17]. For example, we can compute the *perceptual distance* between two colors using metrics from color science like CIEDE2000 [SWD05] or from visualization [Sza18]. Guidelines further note that hue variation and lightness difference specifically impact people's color palette interpretation [TQWS23, KRC02]. As the difference between colors increases, perceptual distance metrics become less useful. However, measures such as color name difference [HS12] provide a metric for color uniqueness or relatedness for larger color differences [GLS16]. Color palettes also must consider people's aesthetic preferences, which, contrary to categorical color design recommendations, may increase with hue similarity [SP11, PSS13]. For example, *pair preference* models human preference using aesthetic color appearance factors including lightness and saturation [SP11].

These metrics inform broader design guidelines for assembling sets of colors into colormaps for representing data. Perceptually-uniform colormaps improved accuracy and response time when comparing values [LH18]. Diverging colormaps support more robust pattern perception than sequential colormaps in heatmaps [RNAK18]. Reda & Szafir [RS20] found that increased color name variability improves graphical inference. Color design tools like Colorgorical [GLS16] leverage these findings using a set of color perception, aesthetic, and naming models to score palettes.

Effective color use goes beyond task alignment. For example, people more effectively interpret visualizations when colors align with their corresponding semantic concepts [SLWF18, KWS19]. Schloss et al. [SGS*18] found that people use implicit color associations to infer values from colormaps, such as dark-is-more or opaque-is-more biases. Follow-up studies suggest semantics are as important as perceptual measures in categorical visualization [MYS*21].

Despite numerous perceptual measures relevant to categorical color palette design, most guidelines for constructing color ramps are loosely-defined qualitative heuristics, posing challenges for users without professional experience [SWS19]. We take a step towards better bridging heuristic and perceptual measures in categorical colormap design by investigating hue, lightness, and uniformity variations across different palette design families.

3. Methodology

We analyzed how the color palettes used to distinguish categories impact people's abilities to reason with multiclass scatterplots. We conducted a crowdsourced study measuring how well people compare category means over varying category numbers ($N = 2 - 10$) and five color palette families (Figure 1). These families systematically vary in hue and lightness, representing common perceptual properties used in palette design. Given the importance of overall perceptual discriminability and uniformity between categories, we additionally tested perceptually-uniform sequential palettes to understand whether uniformity may play a role in categorical encoding, where color differences between marks tend to be larger than in standard unit steps with continuous encodings. We hypothesized:

H1: Categorical palettes will perform better than all other palettes. Categorical palettes are designed with best practices and principles for categorical data encoding. These palettes emphasize hue variations while conforming to effectiveness and expressiveness principles for categorical data [Mun14], unlike diverging or sequential palettes designed to communicate quantitative data.

H2: Diverging palettes will perform better than sequential palettes. Diverging palettes encourage people to cluster data by using hue to indicate the end and middle points of a distribution. This grouping may reduce people's abilities to estimate quantitative statistics within a palette compared to sequential encodings [LH18]; however, this reduction may indicate stronger grouping cues than sequential palettes with similar hue variation which outperform diverging palettes on quantitative estimation tasks.

H3: Multi-hue sequential palettes will perform better than single-hue sequential palettes. Prior experiments reported using multi-hue palettes to represent data could help people distinguish data and find correlations [LH18,SKR99,Lev96,BI07]. Given hue's importance in palette design heuristics, we anticipate palettes varying more in hue will lead to better performance.

H4: Perceptually-uniform sequential palettes will perform better than other sequential palettes. Many popular sequential palettes are perceptually non-uniform [SWS19]. Perceptual distance metrics help align colors with human perception [GPML*16, TQWS23] and have commonly been applied in color measurement and palette design [GLS16,HS12,SWD05]. We anticipate that more uniform steps between colors will improve people's abilities to distinguish colors across palettes by ensuring a sufficient minimum difference between pairs of categories and perceived category differences that are proportional to the number of categories.

3.1. Task & Stimuli

We used a relative mean judgment task from previous studies [GCNF13, SHGF16, HWS21, TQWS23]. As in Tseng et al. [TQWS23], we asked participants to estimate the category with the highest average y-value (e.g., "*Which category has the highest average mean value?*"). This task required participants to first find the data points of each category and then estimate statistical values over all points in that category such that confusion between categories (colors) leads to errors in estimation.

We generated each scatterplot as a $400px \times 400px$ graph using D3. Scatterplots were rendered on a white background using two orthogonal black axes with 13 unlabeled ticks. We generated point positions according to the data distribution methods in Tseng et al. [TQWS23] with a "medium" hardness level and 15 points with a four-pixel radius-filled circle in each category to generate 180 datasets. We selected 20 color palettes for five color families (four per family). Figure 1 shows the palettes and Figure 2 the example stimuli. We chose these color palettes from expert-crafted examples based on their performance in previous studies [SWS19,TQWS23], their high relative lightness magnitude or variance [ZHM09], and high *perceptual distance* between colors [SWD05]. For categorical palettes, colors were randomly selected. We sampled colors from continuous palettes by mapping them to the range [0, 1] and using uniform data steps to choose the corresponding categorical colors.

3.2. Procedure & Participants

Our study involved three phases: (1) informed consent and color-blindness screening, (2) task description and tutorial, and (3) formal study. Participants initially gave consent per IRB guidelines, provided demographics, completed an Ishihara test for colorblindness, and then were introduced to the tasks. Participants clicked on a data point in the category with the highest average y-value in a scatterplot. We randomly divided the 20 color palettes into 4 groups with each group containing one palette from each family. Each participant responded to 45 stimuli containing one combination of each palette from certain group × category number (2–10) presented in random order with a random dataset. We randomly assigned the group when participants started the study. To avoid learning or fatigue effects, three engagement checks were randomly placed along with the formal questions. The engagement checks had three categories with large differences in their means.

We recruited 112 U.S. and Canada-based participants via MTurk. Data from 100 participants (63 male, 37 female, aged 24-65) without CVD and passing all engagement checks were analyzed, with each receiving $3.00 compensation. The anonymized data, stimuli, results, and infrastructure for our study can be found on OSF.

4. Results

Accuracy was the primary dependent measure. To compare performance across palette designs and category numbers, we conducted a two-factor ANOVA with color palette family and number of categories as independent variables and Tukey's HSD with $\alpha = 0.05$ and Bonferroni correction for post-hoc analysis.

Color Families: Our results show that color palette families significantly impact performance ($F(4, 95) = 12.25, p < .0001$). As shown in Figure 1, categorical palettes (91.44% accuracy) significantly outperformed diverging (86.78%), perceptually-uniform sequential multi-hue (86.67%), sequential multi-hue (82.56%), and sequential single-hue palettes (81.11%), supporting **H1**. Diverging and perceptually-uniform multi-hue palettes outperformed sequential single-hue palettes. These findings partially support **H2**: diverging palettes outperformed single-hue sequential palettes.

Our results partially support **H3**: perceptually-uniform multi-hue sequential palettes performed better than single-hue sequential palettes; however, perceptually-uniform multi-hue palettes did not outperform general multi-hue sequential palettes. In partial support of **H4**, perceptually uniform multi-hue sequential palettes performed marginally better than non-uniform multi-hue palettes. However, further analysis revealed that perceptually uniform palettes traversed a larger range of color differences than non-uniform multi-hue palettes. Given the larger range of differences and predictable difference structure imparted by unifor-
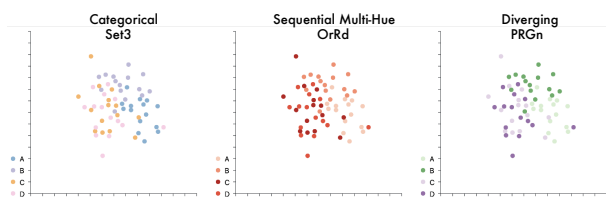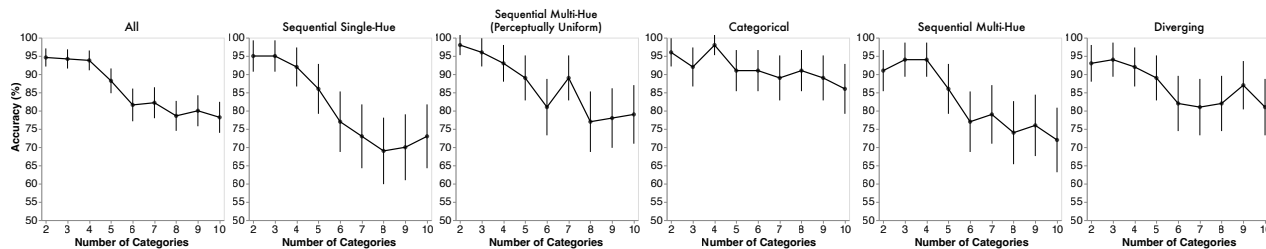

Figure 2: *Example stimuli for three palette types.*

Figure 3: *Accuracy rates broken down by category number separated by color palette families. We found that categorical palettes were more robust to increasing numbers of categories, with other families dropping in performance for six or more categories.*

mity, one might expect perceptually-uniform palettes to dramatically outperform non-uniform palettes, contrary to the relatively small difference in our findings. Part of this difference may pertain to degradation in perceptual distance metrics at larger differences [HS12, GLS16]. Future work should explore the role of perceptual uniformity and related metrics in categorical palette design.

Number of Categories: Our results confirm Tseng et al.'s findings [TQWS23] that *performance decreased as the number of categories increased*. Our results reveal a significant effect of the number of categories on judgment accuracy ($F(8,91) = 20.68, p < .0001$): people were less accurate with higher numbers of categories (Figure 3). For every family, the accuracy rate decreased as the number of categories increased. However, the categorical palettes remained relatively robust even with higher numbers of categories. Sequential and diverging palettes remained notably robust to small numbers of categories (less than six).

5. Discussion

When designing categorical encodings, our results empirically confirm heuristic guidance that hue is the most significant factor. However, other factors also likely play a role in encoding robustness. Key findings and design implications from our study include:

Categorical palettes effectively encode categorical data. Predominantly hue-varying categorical color palettes achieved the highest and most robust accuracy among all palette types. This finding validates heuristic best practices for categorical encoding design. Further, it provides evidence in support of the expressiveness and effectiveness principles [Mun09, QWW*24]: palettes that maximize intuitive categorical attributes outperformed those that maximize ordinal attributes (i.e., sequential and diverging palettes).

Hue is the primary factor for effective palette design, but may be insufficient when considered alone. The average performance and robustness of multi-hue sequential palettes were higher than single-hue. This finding aligns with several existing studies [LH18, SKR99, Lev96, BI07], which indicate multi-hue sequential palettes make it easier to find correlations between certain colors.

However, our results also indicate that hue alone is insufficient to explain performance. For example, mean overall performance and robustness for the single-hue and non-uniform multi-hue color palettes were nearly identical despite multi-hue palettes having greater hue variation, more nameable hues, and larger overall perceptual distance between colors. Diverging palettes performed slightly better overall and were more robust than sequential palettes

with as many or more distinct hues. These results, coupled with findings from previous exploratory studies that indicate the importance of lightness variation in categorical encodings [TQWS23], suggest an interplay between lightness, hue, and categorical perception. These variations may also relate to perceived ordering or similar factors. Understanding these factors and their interactions is important future work.

Increasing category numbers influences performance. Our experiments support that adding more categories significantly reduces categorical perception accuracy [TQWS23]. Further, these findings partially replicate the subitizing phenomenon found in past work by revealing a noticeable performance drop between five and six classes [TQWS23, HW12]. In our case, performance appears to fall rapidly between five and six categories but remains relatively stable for six to ten categories. This stability may indicate a change in the perceptual strategy used to process larger numbers of categories that should be interrogated in future work.

Our results provide mixed evidence of how palette design might influence this threshold. We did not find a significant drop in categorical palettes; however, diverging palettes, which are essentially two single-hue sequential palettes joined by white in our study, exhibited the same "S"-shaped dip in performance at six categories as sequential palettes, with slightly higher average performance after six categories. Exploring this tension and how design might mitigate (or leverage) changing perceptual strategies in categorical data visualization is important future work.

Limitations & Future Work: This study examines design factors in categorical encodings using a small set of expert encodings for mean estimation. While largely consistent with Tseng et al.'s findings [TQWS23], discrepancies in categorical palette performance rankings, possibly due to varied study settings, require further analysis. Contrary to expectations, perceptually-uniform palettes with larger average color step sizes did not significantly outperform non-uniform ones, suggesting a need for more research into the impact of perceptual metrics on palette design. Future work should also explore the potential misinterpretation of category ordering in sequential encodings to more holistically characterize color palette design's role in expressiveness and effectiveness.

6. Acknowledgments

References

[BI07] BORLAND D., II R. M. T.: Rainbow color map (still) considered harmful. *IEEE Comput. Graph. Appl. 27*, 2 (2007), 14–17. 3, 4

[BSG17] BURLINSON D., SUBRAMANIAN K., GOOLKASIAN P.: Open vs. closed shapes: New perceptual categories? *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2017), 574–583. 2

[BTS*17] BUJACK R., TURTON T. L., SAMSEL F., WARE C., ROGERS D. H., AHRENS J.: The good, the bad, and the ugly: A theoretical framework for the assessment of continuous colormaps. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2017), 923–933. 2

[DBH14] DEMIRALP Ç., BERNSTEIN M. S., HEER J.: Learning perceptual kernels for visualization design. *IEEE Trans. Vis. Comput. Graph. 20*, 12 (2014), 1933–1942. 2

[GCNF13] GLEICHER M., CORRELL M., NOTHELFER C., FRANCONERI S.: Perception of average value in multiclass scatterplots. *IEEE Trans. Vis. Comput. Graph. 19* (2013). 1, 2, 3

[GH10] GOLDSTONE R. L., HENDRICKSON A. T.: Categorical perception. *Wiley Interdiscip. Rev.: Cogn. Sci. 1*, 1 (2010), 69–78. 2

[GLS16] GRAMAZIO C. C., LAIDLAW D. H., SCHLOSS K. B.: Colorgorical: Creating discriminable and preferable color palettes for information visualization. *IEEE Trans. Vis. Comput. Graph. 23*, 1 (2016), 521–530. 1, 2, 3, 4

[GPML*16] GÓMEZ-POLO C., MUÑOZ M. P., LUENGO M. C. L., VICENTE P., GALINDO P., CASADO A. M. M.: Comparison of the cielab and ciede2000 color difference formulas. *J. Prosthet. Dent. 115*, 1 (2016), 65–70. 3

[HB03] HARROWER M., BREWER C. A.: Colorbrewer. org: an online tool for selecting colour schemes for maps. *Cartogr. J. 40*, 1 (2003), 27–37. 1

[HS12] HEER J., STONE M.: Color naming models for color selection, image editing and palette design. In *ACM CHI* (2012), pp. 1007–1016. 2, 3, 4

[HW12] HAROZ S., WHITNEY D.: How capacity limits of attention influence information visualization effectiveness. *IEEE Trans. Vis. Comput. Graph. 18*, 12 (2012). 4

[HWS21] HONG M.-H., WITT J. K., SZAFIR D. A.: The weighted average illusion: Biases in perceived mean position in scatterplots. *IEEE Trans. Vis. Comput. Graph. 28*, 1 (2021), 987–997. 1, 3

[HYFC14] HARRISON L., YANG F., FRANCONERI S., CHANG R.: Ranking visualizations of correlation using weber's law. *IEEE Trans. Vis. Comput. Graph. 20*, 12 (2014), 1943–1952. 1

[JQL*23] JEON H., QUADRI G. J., LEE H., ROSEN P., SZAFIR D. A., SEO J.: Clams: A cluster ambiguity measure for estimating perceptual variability in visual clustering. *IEEE Trans. Vis. Comput. Graph.* (2023). 1

[KRC02] KINDLMANN G., REINHARD E., CREEM S.: Face-based luminance matching for perceptual colormap generation. In *IEEE VIS* (2002), pp. 299–306. 2

[KWS19] KINATEDER M., WARREN W. H., SCHLOSS K. B.: What color are emergency exit signs? egress behavior differs from verbal report. *Appl. Ergon. 75* (2019), 155–160. 2

[Lev96] LEVKOWITZ H.: Perceptual steps along color scales. *Int. J. Imaging Syst. Technol. 7*, 2 (1996), 97–101. 3, 4

[LH18] LIU Y., HEER J.: Somewhere over the rainbow: An empirical assessment of quantitative colormaps. In *ACM CHI* (2018), p. 598. 2, 3, 4

[Mac86] MACKINLAY J.: Automating the design of graphical presentations of relational information. *ACM Trans. Graph. 5*, 2 (1986), 110–141. 1

[Mun09] MUNZNER T.: A nested model for visualization design and validation. *IEEE Trans. Vis. Comput. Graph. 15*, 6 (2009), 921–928. 4

[Mun14] MUNZNER T.: *Visualization analysis and design*. CRC press, 2014. 1, 3

[MYS*21] MUKHERJEE K., YIN B., SHERMAN B. E., LESSARD L., SCHLOSS K. B.: Context matters: A theory of semantic discriminability for perceptual encoding systems. *IEEE Trans. Vis. Comput. Graph. 28*, 1 (2021), 697–706. 2

[PSS13] PALMER S. E., SCHLOSS K. B., SAMMARTINO J.: Visual aesthetics and human preference. *Annu. Rev. Psychol. 64* (2013), 77–107. 2

[QNWR22] QUADRI G. J., NIEVES J. A., WIERNIK B. M., ROSEN P.: Automatic scatterplot design optimization for clustering identification. *IEEE Trans. Vis. Comput. Graph.* (2022). 1

[QR20] QUADRI G. J., ROSEN P.: Modeling the influence of visual density on cluster perception in scatterplots using topology. *IEEE Trans. Vis. Comput. Graph. 27*, 2 (2020), 1829–1839. 1

[QR21] QUADRI G. J., ROSEN P.: A survey of perception-based visualization studies by task. *IEEE Trans. Vis. Comput. Graph.* (2021). 1

[Qua21] QUADRI G. J.: *Constructing Frameworks for Task-Optimized Visualizations*. PhD thesis, University of South Florida, 2021. 1

[QWW*24] QUADRI G. J., WANG A. Z., WANG Z., ADORNO J., ROSEN P., SZAFIR D. A.: Do you see what i see? a qualitative study eliciting high-level visualization comprehension. In *ACM CHI* (2024), pp. 1–26. 4

[RNAK18] REDA K., NALAWADE P., ANSAH-KOI K.: Graphical perception of continuous quantitative maps: the effects of spatial frequency and colormap design. In *ACM CHI* (2018), pp. 1–12. 1, 2

[RS20] REDA K., SZAFIR D. A.: Rainbows revisited: Modeling effective colormap design for graphical inference. *IEEE Trans. Vis. Comput. Graph. 27*, 2 (2020), 1032–1042. 2

[SGS*18] SCHLOSS K. B., GRAMAZIO C. C., SILVERMAN A. T., PARKER M. L., WANG A. S.: Mapping color to meaning in colormap data visualizations. *IEEE Trans. Vis. Comput. Graph. 25*, 1 (2018), 810–819. 2

[SHGF16] SZAFIR D. A., HAROZ S., GLEICHER M., FRANCONERI S.: Four types of ensemble coding in data visualizations. *J. Vis. 16*, 5 (2016), 11–11. 1, 3

[SKR99] SPENCE I., KUTLESA N., ROSE D. L.: Using color to code quantity in spatial displays. *J. Exp. Psychol. Appl. 5*, 4 (1999), 393. 3, 4

[SLWF18] SCHLOSS K. B., LESSARD L., WALMSLEY C. S., FOLEY K.: Color inference in visual communication: the meaning of colors in recycling. *Cogn. Res.: Princ. Implic. 3*, 1 (2018), 1–17. 2

[SP11] SCHLOSS K. B., PALMER S. E.: Aesthetic response to color combinations: preference, harmony, and similarity. *Atten. Percept. Psychophys. 73* (2011), 551–571. 2

[SS19] SMART S., SZAFIR D. A.: Measuring the separability of shape, size, and color in scatterplots. In *ACM CHI* (2019), p. 669. 2

[Sto06] STONE M.: Choosing colors for data visualization. *Business Intelligence Network 2* (2006). 1

[SWD05] SHARMA G., WU W., DALAL E. N.: The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Res. Appl. 30*, 1 (2005), 21–30. 2, 3

[SWS19] SMART S., WU K., SZAFIR D. A.: Color crafting: Automating the construction of designer quality color ramps. *IEEE Trans. Vis. Comput. Graph. 26*, 1 (2019), 1215–1225. 2, 3

[Sza18] SZAFIR D. A.: Modeling color difference for visualization design. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2018), 392–401. 2

[TQWS23] TSENG C., QUADRI G. J., WANG Z., SZAFIR D. A.: Measuring categorical perception in color-coded scatterplots. In *ACM CHI* (2023). 1, 2, 3, 4

[ZHM09] ZEILEIS A., HORNIK K., MURRELL P.: Escaping rgbland: Selecting colors for statistical graphics. *Comput. Stat. Data Anal. 53*, 9 (2009), 3259–3270. 3