Mapping Color to Meaning in Colormap Data Visualizations: Supplementary Material

Karen B. Schloss, Connor C. Gramazio, A. Taylor Silverman, Madeline L. Parker, and Audrey S. Wang

1 Legend text position in Experiments 1 and 2

In both experiments, we varied the position of the labels in the legend such that "greater" was higher on half of the trials and "fewer" was higher on the other half of the trials (Figure S1). We also orthogonally varied the orientation of the color scale in the in the legend such that dark was high on half of the trials and light was high on the other half. This resulted in two encoded mappings (dark-more and light-more).

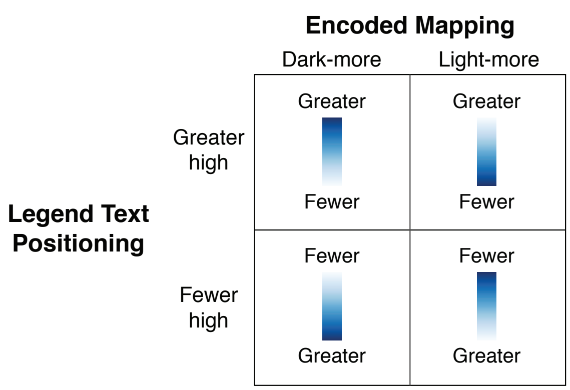

Fig. S1. Four legend conditions: 2 legend lightness mappings (dark-more, light-more) x 2 legend text positionings ("greater"–high, "fewer"–high).

Figure S2 shows mean RTs from Experiment 1, separated by legend text positioning. The overall ANOVA reported in the main text revealed a main effect of legend text positioning. RTs were faster when "greater" was high in the legend than when "fewer" was high in the legend ($F(1,29)$ = 112.25, p <.001, η_p^2 = .795). This result is consistent with prior reports that larger values should be higher in visualizations [1] because people have an inferred mapping between greater and higher. There was also an interaction between legend text positioning and encoded mapping ($F(1,29)$ = 6.64, p = .015, η_p^2 = .186). This interaction revealed that the degree to which RTs were faster for dark-more encoding than light-more encoding was amplified when "greater" was at the top of the legend.

Figure S3 shows mean RTs from Experiment 2, separated by legend text positioning. Again, the overall ANOVA reported in the main text revealed a main effect of legend text positioning. RTs were faster when "greater" was high in the legend than when "fewer" was high ($F(1,29)$ = 53.88, p <.001, η_p^2 = .650). As in Experiment 1, there was an interaction between legend text positioning and encoded mapping ($F(1,29)$ = 5.12, p = .031, η_p^2 = .150), with more extreme differences between dark-more and light-more encoding when "greater" was high in the legend. There was also a 3-way interaction between encoded mapping $\times$ legend text positioning $\times$ background ($F(2,58)$ = 5.38, p = .007, η_p^2 = .157). This was due to the relative differences between encoded mappings on different backgrounds being more extreme when"greater" was high in the legend. For example, for the black–white colormaps, the degree to which RTs were faster for dark-more encoding on light background and light-more encoding on dark backgrounds was larger when "greater" was high than when it was low in the legend.

A potential explanation for these interactions is that there is a ceiling effect when "fewer" is higher in the legend. That is, the processing cost for interpreting legends in which "fewer" is high drowns out any effects due to the encoded lightness mappings.

2 Jet color scale in Experiment 1

In our discussion of the Experiment 1 results in the main text, we focused on the four color scales that vary monotonically in lightness. Here we report on the Jet color scale, which varies non-monotonically in lightness—it is dark at its two endpoints and lightest in the middle (Figure S4A). Therefore, instead of coding the legend conditions according to lightness encoding (dark-more vs. light-more), we coded them according to cool/warm encoding (cool-more vs. warm-more).

The mean RTs for the Jet are shown in Figure S4B. We analyzed the data using a 3-way repeated measures ANOVA: 2 warm/cool encodings (cool-more, warm-more) x 2 backgrounds (white, black) x 2 legend text positions ("greater"–high, "fewer"–high). Overall, RTs were faster for cool-more mapping than warm-more mapping ($F(1,29)$ = 8.579, p = .007, η_p^2 = .228) and this effect did not interact with the background (F<1). Similar to the results reported above, RTs were faster when "greater" was high in the legend than when "fewer" was high ($F(1,29)$ = 44.53, p <.001, η_p^2 = .606), though this factor did not significantly interact with warm/cool encoding ($F(1,29)$ = 3.836, p = .060, η_p^2 = .117).

One interpretation of the warm/cool encoding results is that participants infer that cooler colors map to larger quantities. However, there is an alternative account. The cooler half of the Jet color scale is darker than the warmer half—not only is the blue endpoint darker than the warm endpoint, but also the lightest region of Jet, the yellow section, is biased toward the warmer side. Therefore, the seemingly cool-is-more bias may simply be the dark-is-more bias. To test this hypothesis, it would be necessary to conduct an analogous experiment with color scales that decorrelate lightness and coolness.

3 Color scales in Experiment 1

In the main text we illustrate the color scales from both experiments plotted on the L* b* plane of CIELAB space. We do so to illustrate the degree to which the different color scales deviate from a linear interpolation between the highest contrast color and the background. This 2D view of CIELAB space is adequate for illustrating this point for most color scales and backgrounds. An exception is Autumn on the black background, where it appears that there is little deviance because the deviance is in the a-axis that is compressed in the L* b* plane view. Therefore, we reproduce the L* b* plane in Figure 5 and such that it can be directly compared to the a view of the L* a* plane.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

References

[1] B. Tversky. Visualizing thought. *Topics in Cognitive Science*, pp. 499 – 535, 2011.

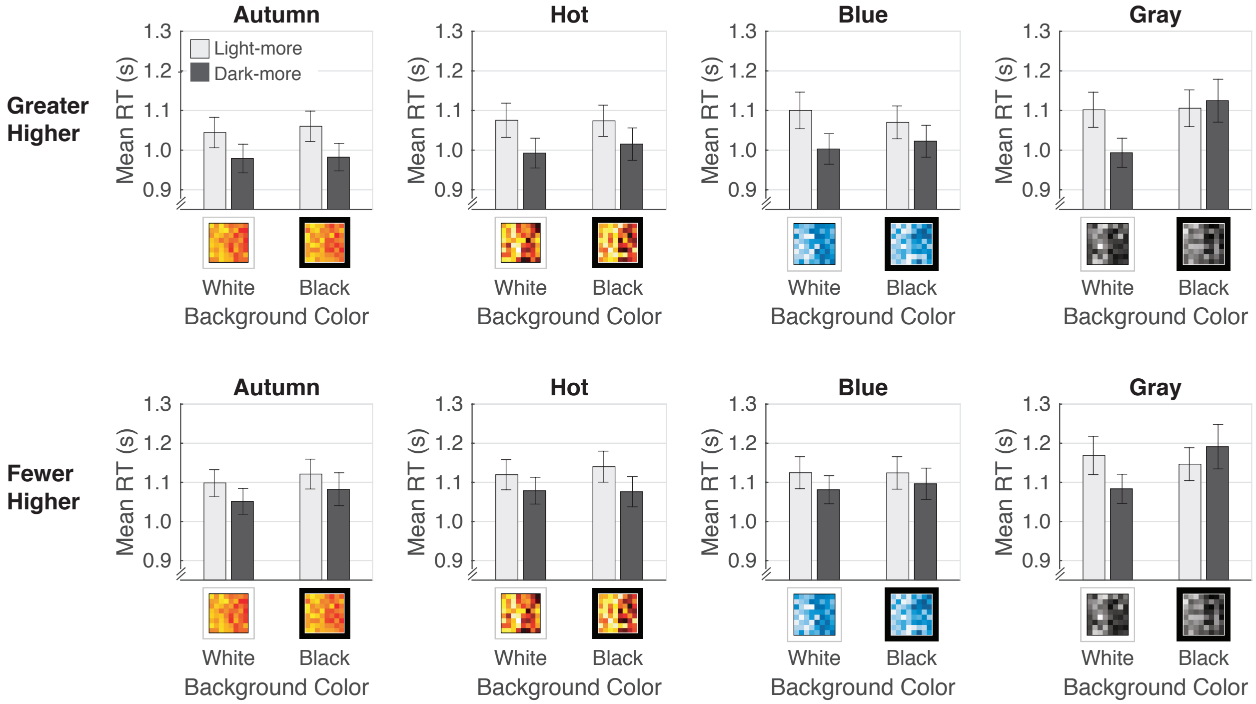

Fig. S2. Mean RTs from Experiment 2, separated by legend text condition: "greater"–high (top row) and "fewer"–high (bottom row), light-more encoding (light bars) and dark-more encoding (dark bars), background color (x-axis), and color scale (separate plots). The icons along x-axis represent example colormaps with each condition (note: the darker region is on the right in these examples, but the dark region was left/right balanced in the experiment). Error bars represent +/– standard errors of the means.

.

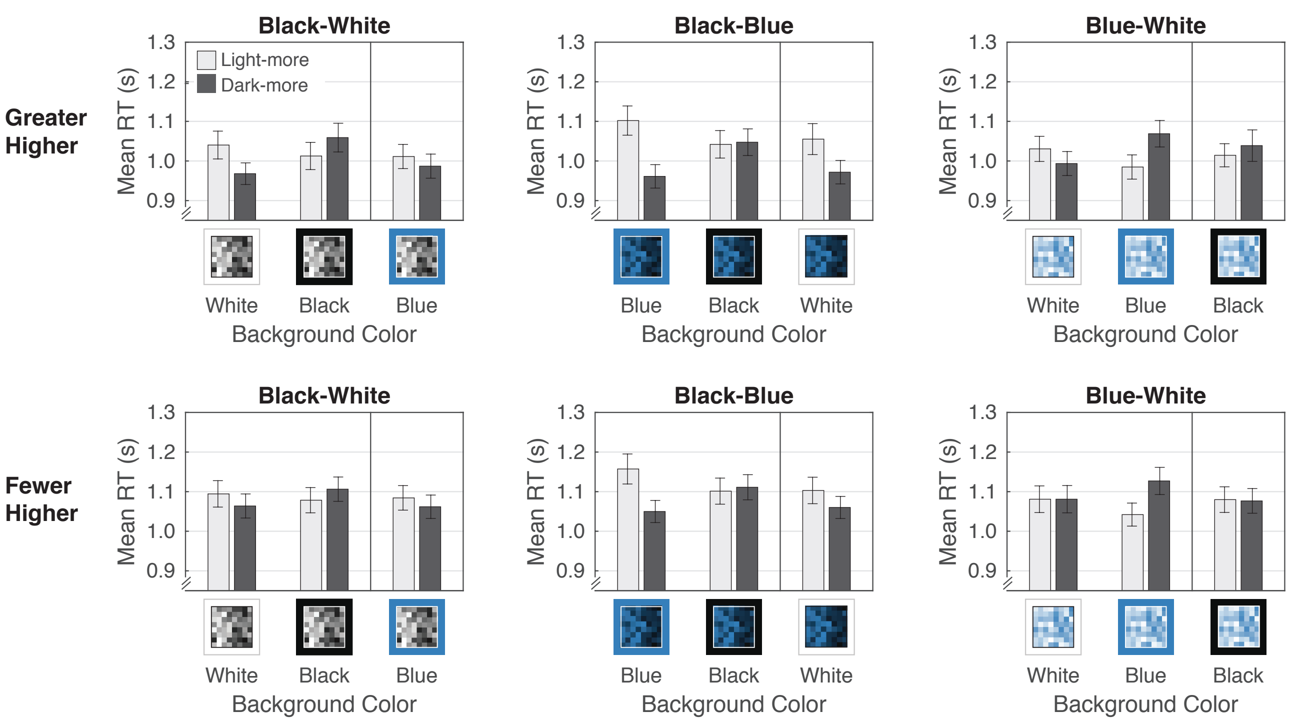

Fig. S3. Mean RTs from Experiment 2, separated by legend text positioning: "greater"–high (top row) and "fewer"–high (bottom row), light-more encoding (light bars) and dark-more encoding (dark bars), background color (x-axis), and and color scale (separate plots). The icons along x-axis represent example colormaps with each condition (note: the darker region is on the right in these examples, but the dark region was left/right balanced in the experiment). The order of the icons along the x-axis is: opacity variation–light background (left), opacity variation–dark background (center), no opacity variation (right). Error bars represent +/– standard errors of the means.

.

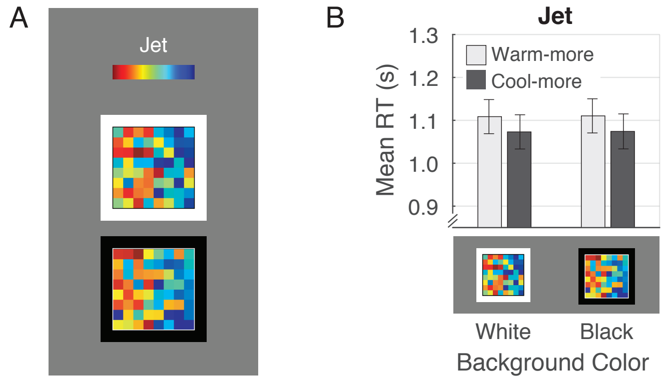

Fig. S4. (A) Jet color scale and corresponding Jet colormaps on a white and black background. (B) Mean RTs for warm-more mapping (light bars) and cool-more mapping (dark bars), separated by background color (x-axis). The icons along x-axis represent example colormaps with each condition (note: the cooler region is on the right in these examples, but the cooler region was left/right balanced in the experiment). Error bars represent +/– standard errors of the means.

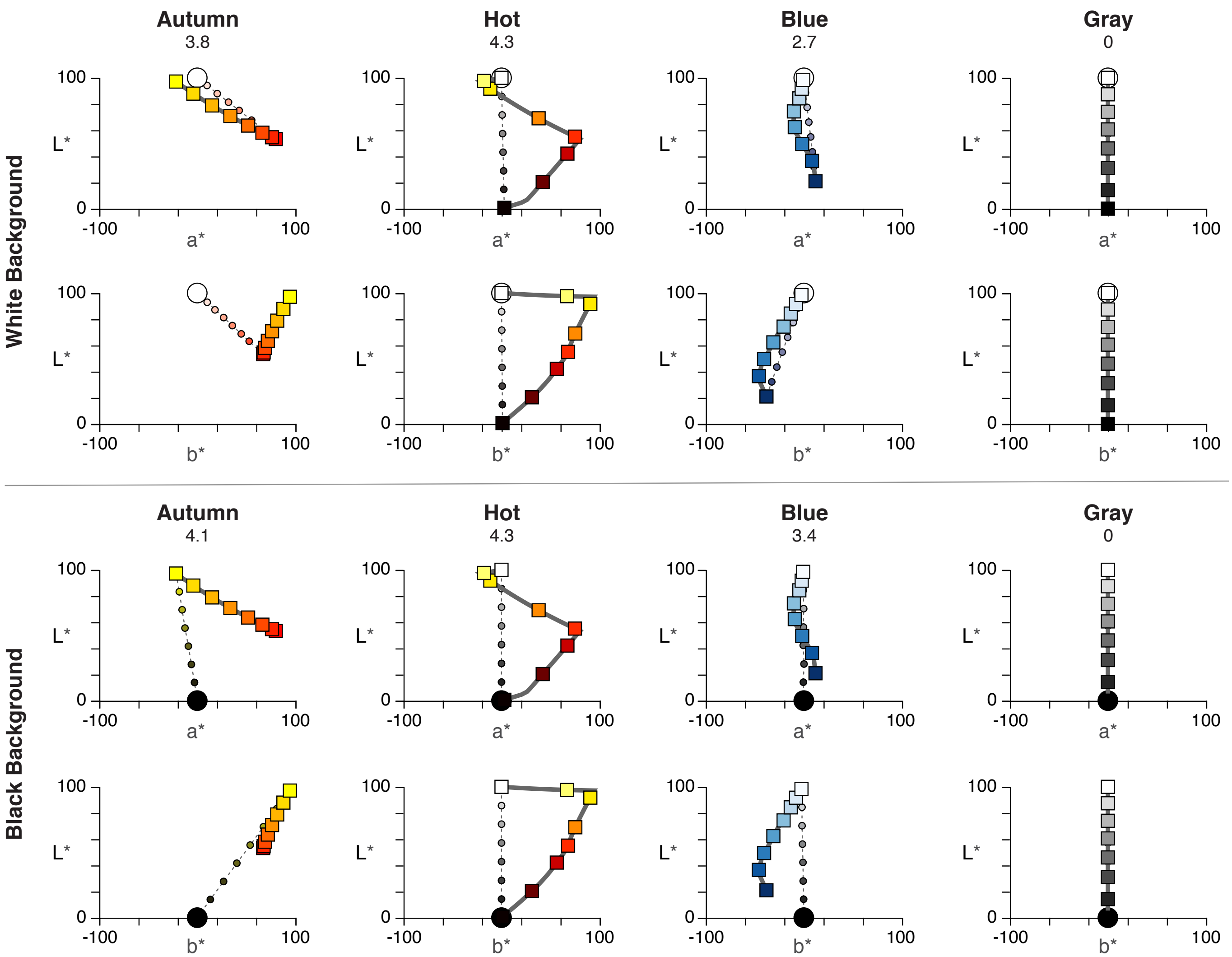

Fig. S5. Color scales from Experiment 1 (squares and thick curve connecting them) and interpolations between the highest-contrast end point for the white and black backgrounds (circles and dashed line connecting them). The colors are ploted in CIELAB space, viewed from the L* a* plane and the L* b* plane. The number below each color scale name is the Opacity Variation Index for that color scale on the white background (top two rows) and black background (bottom two rows).