Weaving versus Blending: a quantitative assessment of the information-carrying capacities of two alternative methods for conveying multivariate data with color

Motivation
Numerous applications require an integrated understanding of the values of multiple scalar variables defined at densely sampled locations over a common domain. For example, in fluid flow research, scientists may want to be able to understand the distributions of velocity magnitude, vorticity magnitude, swirl strength and Reynolds shear stress, or in a mapping application analysts may want to be able to see the distributions of – and gain insight into the relationships between – quantities such as median family income, educational attainment, home ownership rates and population density. As such, the most common approaches historically used to visualize such data are: 1) to separately encode the values of each variable, via a color scale that continuously varies in luminance and/or saturation level in individual images that are either: simultaneously displayed side by side, sequentially displayed in the same location and interactively shuffled, or made semi-transparent and laid on top of one another; or 2) to encode the values of all of the multiple variables at each point into a single color palette pattern that is then displayed over a small region of the domain around each sample point.

In this paper, we present the results of a series of experiments that seek to quantitatively assess the relative effectiveness of two of these representational approaches: color weaving (the most common default approach), in which a single composite color is used to convey the values of multiple color-encoded quantities; and color weaving (a recently developed approach), in which the individual colors of multiple variables are sequentially woven into a fine-grained texture pattern.

Experiment 0 - Goals and Methods
The goal of our first experiment was to assess participants’ baseline ability to accurately read numerical data encoded via the immensity of a single displayed color. We created six different perceptually linear single-band color ramps, defined by continuous variations in the luminance and saturation values of each of six different base hues. We defined the base hue by choosing six evenly-spaced points around a circle of constant luminance in a region of the La*b* color space that fit within our monitor gamut.

We recruited nine participants with normal color vision for this baseline experiment. Participants’ color vision ability was assessed using an online version of a collection of Ishihara plates, the validity of a subset of which had been informally assessed for us by a person with known red-green color perception deficiencies.

The stimuli consisted of six maps of the twelve midwestern United States, in which each state was filled with a different constant color, from a single color ramp, representing the values of a particular data attribute for that state. The actual data values were obtained from 2000 census data, but the particular assignment of data values to states was randomized to prevent people from using domain knowledge about the midwestern US to increase the accuracy of their guesses. The six data distributions were: median household income, percentage of the population that had graduated from high school, percentage of the population that had graduated from college, percentage of the population living below the poverty line, median cost of a single family dwelling, and home ownership rate. On each trial in the experiment, the participant’s task was to identify the value of a particular data attribute for a particular state by reading the color from the map, setting a slider to the matching color on the provided color scale, and then clicking on the state to indicate that their selection is final. Figure 1 (top) shows a screen shot from one trial. All participants were provided with a printed outline map showing the correspondence of names to states, and all trials on which the name was mis-identified were discarded.

Experiment 1 - Results
Participants were uniformly able to perform the assignment with fairly good accuracy. The average relative error, computed over all participants and all colors, was 6.02%, with a standard error of 0.57%. Figure 1 (right) shows a scatter plot of the average errors for each trial, computed over the nine participants. The error bars show the amounts of the 90% confidence intervals, and each point is color-coded according to the base color of the color scale used for that trial.

Experiment 2 - Goals and Methods
Because the color values that we used in experiment 1 were defined according to actual data values, the distribution of measurements that we collected - while highly representative of what would be encountered in a typical visualization application - was not sufficient to answer all of the questions that came up about people’s ability (or inability) to accurately decompose a composite color into its constituent components. Our main goals in experiment 2 were to specifically explore the relationship between hue separation, luminance difference, and the error rates in people’s judgments of the values of individual components in its two-stimuli representation. Specifically, we sought to answer the two questions: 1) will error rates be greater, smaller, or the same for the reading of color combinations in which the hues are separated by 60º, 120º, or 180º in La*b* space?; and 2) will error rates be greater, smaller, or the same for the reading of color combinations in which the luminance values of the individual components are nearly equal, moderately close, or relatively widely separated? The figures below show some color pairs and noise patterns from among the 360 sample stimuli that we used in this final experiment. Four participants provided data for our analysis.

Experiment 2 - Results
In a 3-way ANOVA analysis of the effects of mixture type, hue difference and luminance difference, we found significant main effects of both mixture type and luminance difference, with errors being higher when colors were combined by blending than when they were intermixed, and with the errors being smaller when the luminance values of the component colors were more similar. Looking separately at individual subsets of the data, defined by blend type and hue angle, we found a significant main effect of luminance difference in all cases (marked in red where statistically significant) except when the hues were directly complementary. Rather than having the hue angle have a significant effect on error rates in any subset of the data, the only effect that we found that was significant was the interaction of blend type, hue angle and luminance difference.

Discussion
The results of our three experiments indicate that color weaving is consistently more effective than color blending for conveying the values of individual data distributions and 4 different colors and only begin to rise to a statistically significant extent when the number of component colors increases to six. This advantage exists despite the potential of complications due to simultaneous contrast effects, and persists even when the area subtended by each patch of continuous color is very small.

Although the problem of inferring the values of the component colors in a blended mixture is relatively ill-defined, observers are able to perform this task fairly accurately, within a moderately constrained domain, when presented with pairs of component colors that have nearly equal luminance values, although errors rise as the luminance values of the component colors begin to differ.

We found no significant advantages, in either color blending or color weaving, to using color scales based on component hues that are more widely separated in La*b* color space. On the contrary, we found some indications that extra difficulties may arise when opponent hues are employed.

This research was supported by a grant from the National Science Foundation (CNS-0323471)