Color, Change, and Control for Quantitative Data Display

Penny Rheingans
Department of Computer Science
University of North Carolina
Chapel Hill, NC 27599-3175

Abstract

Color is used widely and reliably to display the value of a single scalar variable. It is more rarely, and far less reliably, used to display multivariate data. Calico, a dynamic tool for the creation and manipulation of color mappings, adds the element of dynamic control over the color mapping to that of color itself for the more effective display and exploration of multivariate spatial data. Using Calico, a one- or two-variable color mapping can be created using parametric equations and a variety of color models. This mapping can be manipulated by moving input devices referenced in the parametric expressions, by applying affine transforms, or by performing free-form deformations. As the user changes the mapping, an image showing the data displayed using the current mapping is updated in real time, along with the images describing the mapping. This paper presents an empirical study investigating the effects of user control and smooth change in the display of quantitative data on user accuracy, confidence, and preference. While the user control did not quite produce a significant effect on accuracy, it did produce significant increases in user preference and confidence.

1. Introduction

Commonly, a researcher wishes to explore a large set of data in order to develop an understanding of the structure and relationships within the data. She may have informal or incomplete hypotheses about that data that she wishes to develop further. This sort of exploratory process differs from more formal hypothesis testing in that specific beliefs about the precise meaning of the data have not yet been formed. Representing the data visually for this exploratory process is appealing because it allows viewers to harness the powerful processing capabilities of the human visual system. Some structures in the data, especially those involving complex spatial relationships and patterns, are easy to detect visually, but difficult to specify for computational detection.

This work is intended to aid researchers whose primary interest is in the spatial structure of the variables under study, in particular, in the spatial structure of the interrelationships between the variables. A researcher interested in understanding the spatial distribution and pattern of a data set might explore whether two variables seemed to be related over a data space, how the geometry of the data space affects such a relationship, and whether isovalue points form some sort of structure. Dynamic representations enable the researcher to explore the temporal consistency of a pattern over a manipulation, providing more insight into the nature of the spatial distribution of the variables. Where spatial correspondence between the variables is important, a representation with both variables displayed in the same image is preferable to a representation where each variable is displayed in its own image. Displaying both variables on the same map can show subtle differences between spatial patterns such as spatial offsets or shape differences.

Representing multivariate data using color is attractive because the human visual system is capable of differentiating easily among hundreds of colors. Data values can also be displayed using color in a smaller area than they could be using parameters such as texture or shape. Selecting a one-variable color mapping which shows, or can be manipulated to show, the data effectively can be a difficult task. Selecting an effective two-variable color mapping is even more difficult. Although some study has been given to evaluating the effectiveness of one variable-color or discrete two-variable mappings [for example, Pizer and Zimmerman 83; Ware 88; Levkowitz 88; Wainer and Francolini 80], almost none has been given to the experimental evaluation of continuous two-variable mappings. This sort of experimental evaluation is an important part of research into the effective display of multivariate data. This paper presents a tool for the selection and manipulation of two-variable color mappings along with an experimental evaluation of a family such mappings.

2. Related Work

A number of systems have been developed for the interactive design of one-variable color mappings. Some of these systems appear as parts of commercially available visualization packages; a few which have appeared in the literature are described below. Robertson built a system which interactively displays color gamuts of display devices to help data analysts visualize perceptual color spaces and understand their components [Robertson 88]. Cox developed a tool for the interactive design of single variable color mappings using spatially periodic, functional control over the red, green, and

0-8186-2897-9/92 $03.00 © 1992 IEEE

blue components [Cox 88]. Guitard and Ware built a similar tool where color sequences were described by curves for their hue, saturation, and value components [Guitard and Ware 90]. An earlier version of Calico allowed for the dynamic manipulation of one-variable color mappings using either RGB or HLS space [Rheingans and Tebbs 90]. None of these systems allowed for the design of two-variable color mappings and none seem to have been evaluated experimentally.

Other researchers have established a theoretical basis for the design of bivariate color sequences. Trumbo [81] presented a conceptual framework for the selection and evaluation of bivariate color sequences. He proposed four basic criteria for effective displays. *Order* : if data values are ordered, the colors chosen to represent them should be perceived as ordered. *Separation* : significantly different levels of variables should be represented by distinguishable colors. *Rows and Columns* : if preservation of univariate information is important, then display parameters should not obscure one another. *Diagonal* : if detection of positive association of variables is a goal, then the displayed colors should be easily identified as belonging to one of three classes : those near the major diagonal, those above it, and those below. Trumbo uses these principles to develop a number of color sequences.

Robertson and O'Callaghan [86] address the problem of realizing Trumbo's color sequences in a perceptually uniform color space. In a perceptually uniform color space, equal Euclidean distances within the space correspond to equal perceptual differences. The LUV and LAB spaces developed by CIE are examples of perceptually uniform color spaces.

Pham [90] proposed the use of spline-based curve, surface, and subspace generation to implement the concepts of Trumbo [81], Robertson and O'Callaghan [86], and Ware [88] for the design of color sequences for univariate, bivariate, and trivariate data mappings. B-splines provided a mechanism for the interactive creation of color sequences through a set of key colors and local sequence modifications. Although sequences were created interactively, no immediate display of data using those sequences seems to have been implemented.

3. Color Mappings

A mapping from data variable values to display colors is composed of a color space which defines the components of color and a curve or surface through that space. A *color space* is the space spanned by three orthogonal basis vectors. Each basis vector corresponds to one component of a color model. For example, the RGB color space is cube-shaped and the basis vectors correspond to the red, green, and blue components.

The *color path* is a one-dimensional parametric curve in the color space which defines the color gamut of the mapping. As the path curves through the color space it completely describes the sequence of colors used in a mapping from a set of values of a single scalar variable to a set of colors (**f**(s), s = distance along the path). For example, if median family income for U.S. counties is mapped to a combination of hue and lightness using a rainbow scale in the HLS model, the color path runs through the hues in an ascending spiral from black to white. See Figure 1. Counties with a low median family income would be displayed in dark reds, those with an average median income in medium greens and blues, and those with a very high median income in pale purples.

The color gamut for a mapping from the values of two scalar variables to a single color is described by a surface through the color space. At each point, this *color sheet* shows the color used to represent a particular combination of the values of the two data variables. When all combinations of values are considered, a surface is formed. For example, a color scheme might map mean education level to hue and median income to brightness (using an HLS space). See Figure 2. Areas with low education levels would be reds, dark when median income is low and pale when it is high. Areas with a relatively average education level would be blues, dark when median income is low and pale when it is high.

4. Calico

In order to create useful color representations, a user must have a view of the entire gamut of representation colors and a view of the data displayed using this mapping. Additionally, a view of the color gamut in context, for example within a color space, can help the user better understand the source of the mapping. If the user can also manipulate the mapping directly and in real-time, he or she can more easily explore the data by watching how the data image changes as the color mapping is changed.

In Calico, the color space appears in the center of the screen, with the color gamut (path for one variable or sheet for two variables) represented by a curve or surface within the space. Figure 2 shows the Calico display for a mapping of two scalar variables. The color gamut appears in the lower right of the screen. The upper left portion of the screen contains the image space showing how a 2D data set is represented using the current color scheme. Changes to the mapping are made in the color space and reflected in real-time in the image space.

A color space is represented by colored samples scattered through the portion of the space which is displayable on the current display. These samples are regularly spaced in the underlying coordinate system of the space. In Calico, color mappings can be created, viewed, and manipulated using any one of the RGB, Smith's HLS [Smith 78], or CIE LUV color models.

In Calico, color paths and sheets are implemented using splines through a set of control points. Color sequence control points are generated from parametric expressions which specify the color component values in terms of data variable values and input device positions. The parametric expressions can contain symbolic data variables (u,v), dynamic input device variables (S for slider position; TX, TY, and TZ for joystick deflection), arithmetic operators, and built-in function calls. The color mapping shown in Figure 2 would be specified by the equations :

Hue = u
Lightness = v
Saturation = 1.0

In the same example, the saturation can be tied to a slider by changing the third equation to :

Saturation = S

Now when the slider is moved to its maximum position, the representation shows saturated colors of varying brightness. When the slider is moved to its minimum value, the saturation is reduced to zero and the representation reduces to a grey scale showing median income as lightness; no information about education level is visible in the example image. As the slider is moved slowly up from minimum, the hues representing education level gradually fade back in

Expressions containing input device variables are re-evaluated when the corresponding input device is moved. A color sheet can be edited by affine transformations or by grabbing and dragging control points of the sheet. During such freeform deformations, a slider controls the degree to which nearby control points are dragged along with the grabbed point. Both the example image and the geometry of the color sequence change in real-time as the user manipulates the input devices.

5. Experimental Design

An empirical study was conducted to investigate the roles of user control of the color mapping and smooth change between images in the exploration of quantitative multivariate data. Specifically, it was hypothesized that dynamic control over the representation would increase subject accuracy and confidence in a data comprehension task. Additionally, it was expected that smooth change between representations would also increase subject accuracy and confidence. In order to make training and testing times manageable, subjects were limited to a single type of manipulation of a particular family of color mappings.

Design. This experiment employed a two-factor, within-subject, partially-counterbalanced design. This design balanced the effects of both trial order and data sets. The two factors determining the type of representation were the degree of control over one representation parameter (the balance between the two colors representing the two data variables) and the smoothness of change between levels of that parameter. Three levels of the control factor and two levels of the smoothness factor were presented. See Figure 3. Representations provided either no control over the balance parameter, control only over pacing of different views, or complete control over the balance between the two colors. Additionally, representations displayed either discontinuous transitions between parameter levels or relatively smooth change between levels (approximately 10 frames per second of user motion).

Display. Stimulus images were displayed on a 512 X 512 Tektronix model 690SR monitor driven by the Pixel-Planes 4 graphics system [Fuchs, et. al. 85]. See Figure 4. A thematic map of the continental United States occupied the upper left quarter of the screen.The thematic map showed two socio-economic variables collected by the 1980 Census. Each county was colored to display the value of the variables using a particular representation scheme. A legend showing the range of colors used to represent each data variable occupied the lower right section of the screen. The center of the screen contained a wire-frame sheet showing the location of the color gamut in the HLS color space. The color space and sheet were turned off if the subject wished. Generally, a subject requested that the space and/or sheet be turned off for either all or none of the trials. Since a within-subject design compares a subject's score under one test condition to that subject's score under other test conditions, this should have little or no effect on the experimental results.

Thematic maps were displayed using one of the six representations shown in Figure 3. In each representation, one data variable was mapped to levels of green and the other to levels of purple. See Figure 4. This family of mappings was chosen because it satisfies Trumbo's criteria of order, separability, preservation of univariate information, and diagonality. Specifically, the use of complementary colors as the display parameters ensures the displayed colors resolve into three basic classes : roughly equal magnitudes of the two variables (greys), $u > v$ (purples), and $v > u$ (greens). In each trial, many images were shown, each with a different balance between the relative contributions of the two variables, ranging from only one variable to only the other. Figure 5

		Control		
		None	Pace	Complete
Change	Jerky	Slide Show	Slide Projector	Interactive
	Smooth	Constant Loop	Multispeed Loop	Dynamic

Figure 3. Experimental Variables and Representations.

shows three images of the data, each with a different balance between the relative contributions of the two variables. The representations differed in how the subject controlled the relative contributions and how often updates occur. The six representations were :

1. Slide Show Representation : Subjects viewed multiple static images of the data, with varying relative contributions of the two parameters. Five unique images were shown in a repeating loop. A new view appeared every 5 seconds. The subject had no control over the content or pacing of the images.
2. Slide Projector Representation : Subjects viewed multiple static images of the data, with varying relative contributions of the two parameters. Five unique images were shown in a repeating loop. A new view appeared when the subject pressed a button. The subject had control over the pacing, but not the content, of the images.
3. Interactive Representation : Subjects viewed multiple static images. Subjects interactively manipulated the representation by selecting values for the balance between the two parameters using a slider valuator and pressing a button to generate the new representation. The subject had control over both the pacing and content of the images.
4. Constant Loop : Subjects viewed a single precomputed film loop. This film loop showed the effects of smoothly varying the relative contributions of the two parameters, but did not allow the subject to control the manipulation. The loop contained 34 unique images with a full cycle completing every 3.5 seconds.
5. Multispeed Loop : Subjects viewed a single precomputed film loop showing the effects of smoothly varying the relative contributions of the two parameters. Subjects controlled the speed of the loop using a slider valuator. The speed selected ranged from full stop to a complete cycle each half second.
6. Dynamic Manipulation : Subjects dynamically manipulated the relative contributions with a slider valuator. The displayed image changed dynamically in response to these manipulations.

Subjects. The 12 subjects were volunteers recruited from among the graduate students and staff of the UNC Computer Science Department. All subjects were found to have normal color vision by the North Carolina Department of Motor Vehicles.

Procedure. Each subject participated in two sessions. The first session consisted of an introduction to the representations and their manipulation followed by six trials, one trial using each representation. The introduction consisted of a written, tutorial-like presentation of each type of representation for data similar to, but not the same as, that used in the rest of the experiment. In each trial, subjects were given a written description of the data and representation for that trial and asked to explore the data set while filling out a worksheet of questions about the data. The worksheet in each trial was the same except for references to the particular variables presented in that trial. See the appendix for an example worksheet. After all trials were complete the subject filled out a final questionaire comparing the representations. The second session consisted of six additional trials, one trial using each representation, followed by another final questionaire.

6. Results

In all analyses, a within-subjects analysis was used; that is, the scores of a subject using one representation method were compared with the scores of that same subject using other representation methods. The analyses performed were comparison of percentage error difference between representations for a single-variable question, comparison of percentage error difference between representations for a two-variable question, two factor analysis of variance (ANOVA) of accuracy attributable to manipulability and smooth change of a representation, and two factor ANOVA of confidence data. Additionally, subjects' preferences for representations were examined. In the sections below, mean values for preference, percent error, and confidence are given to convey a sense of the direction and pattern of differences.

Subject Preferences. Subjects were asked to rank the representations from 1 to 6 (1 is most preferred; 6 is least preferred). Almost without exception, subjects ranked the dynamic representation as the most preferred, usually followed by the interactive representation. Subjects almost always ranked either the slide show or constant loop as the least preferred representation. Representations were usually ranked in order of increasing control. Within a pair of representations with the same amount of control, for example Slide Projector and Multispeed Loop, some subjects usually ranked the representation with smooth change higher, while other subjects usually ranked it lower. Figure 6 shows the pattern of preference means. Most subjects gave identical rankings after the two sessions. There was no clear pattern of change among those subject who changed their rating between the sessions.

In comments about why they ranked the representations as they did, most subjects mentioned that they liked having control over the representation. Accordingly, many subjects found the interactive representation to be almost as good as the dynamic. Some subjects also mentioned liking smooth change between parameter balance levels. On the negative side, many subjects mentioned that they were most frustrated by representations where they had to wait for the image they wanted or where they could not freeze the display on a particular view.

Error Differences. For each data set, subjects answered questions about the value of a variable in an area (one-variable question — see question 1 on the sample worksheet in the appendix) or about the value of a variable in areas where the value of the other variable met some criterion (two-variable question — see question 2 on the sample worksheet). On one-variable questions, there were small, almost significant ($p < 0.10$ using Student's t-test) differences in error rates between representations. Figure 7 shows the pattern of means. Figure 8 shows the analysis of variance. On this question, subjects had lower error rates using manipulable representations than using nonmanipulable representations. Subjects also had slightly lower error rates using representations which did not provide smooth change. There was no real coherent pattern of differences in accuracy on the two-variable question. This does not necessarily mean that there was no difference in accuracy, but does mean that any such difference was dwarfed by the variability in error rates.

		Control		
		None	Pace	Complete
Change	Jerky	Slide Show 5.12	Slide Projector 4.00	Interactive 2.71
	Smooth	Constant Loop 5.04	Multispeed Loop 2.92	Dynamic 1.08

Figure 6. Pattern of means : representation preferences. A score of 1 shows the most preferred representation, 6 shows the least preferred.

		Control		
		None	Pace	Complete
Change	Jerky	Slide Show 11.9	Slide Projector 8.0	Interactive 8.1
	Smooth	Constant Loop 12.0	Multispeed Loop 9.7	Dynamic 9.1

Change	Jerky	Slide Show 11.0	Slide Projector 9.8	Interactive 10.8
	Smooth	Constant Loop 10.6	Multispeed Loop 11.9	Dynamic 12.4

Figure 7. Pattern of means : percent error. Error means for each representation as a percentage of the range of that data variable. The upper table shows error means for one-variable questions; the lower table shows error means for two-variable questions.

Confidence. Subjects were asked to rate their confidence in their answers on a scale from 1 to 10. In general, subjects were significantly more confident ($p < 0.01$ and $p < 0.025$, respectively for the one- and two-variable questions) about their answers using manipulable than using nonmanipulable representations. This means that there is less than a 1 or 2.5 percent chance that the observed difference is the result of normal variation. The two-factor ANOVA also shows an effect of the interaction between degree of control and smoothness of change that approaches significance. Specifically, control was even more important in representations that were changing smoothly. Figure 9 shows the pattern of means while Figures 10 and 11 show the analysis of variance.

7. Discussion

Results suggest that subjects used their control over the representations partly to remove unwanted information. This was most apparent on the single-variable question where subjects manipulated the representation to show only the desired variable. Accordingly, the answers to this question were slightly more accurate when the subject could manipulate the representation. A few subjects manipulated the representation this same way when answering the two-variable question, first viewing one variable, then the other, with few or no intermediate steps. This strategy was less successful on this question, resulting in answers that were no more accurate than using nonmanipulable representations.

Subjects found that different representations and manipulations were best suited for answering different types of questions. One subject commented that the interactive and dynamic representations were good for singling out one variable, whereas the dynamic and loop representations were good for answering two-variable questions. Presumably this

source	SS	df	MS	F	
subjects	0.04	11	0.00		
control	0.02	2	0.01	2.60	$p < 0.10$
dynamic	0.00	1	0.00	0.35	
CxD	0.00	2	0.00	0.06	
CxS	0.07	22	0.00		
DxS	0.05	11	0.00		
CxDxS	0.15	22	0.01		
Total	0.33	71			

Figure 8. Two-factor ANOVA for one-variable accuracy.

		Control		
		None	Pace	Complete
Change	Jerky	Slide Show 7.63	Slide Projector 8.08	Interactive 8.04
	Smooth	Constant Loop 7.25	Multispeed Loop 8.21	Dynamic 8.42

Change	Jerky	Slide Show 5.42	Slide Projector 6.21	Interactive 5.92
	Smooth	Constant Loop 5.29	Multispeed Loop 6.00	Dynamic 6.75

Figure 9. Pattern of means : confidence. Subject confidence in responses on a scale from 1 to 10. The upper table on shows confidence levels for one-variable questions; the lower table shows confidence levels for two-variable questions. In both tables, 1 shows minimal confidence, whereas 10 shows extreme confidence.

source	SS	df	MS	F	
subjects	34.34	11	3.12		
control	9.08	2	4.54	6.38	$p < 0.01$
dynamic	0.03	1	0.03	0.07	
CxD	1.75	2	0.88	1.20	
CxS	15.67	22	0.71		
DxS	4.59	11	0.42		
CxDxS	16.00	22	0.73		
Total	81.47	71			

Figure 10. Two-factor ANOVA for one-variable confidence.

source	SS	df	MS	F	
subjects	56.68	11	5.15		
control	13.00	2	6.50	4.41	$p < 0.025$
dynamic	0.42	1	0.42	0.32	
CxD	3.69	2	1.85	1.57	$p < 0.25$
CxS	32.42	22	1.47		
DxS	14.37	11	1.31		
CxDxS	25.89	22	1.18		
Total	146.47	71			

Figure 11. Two-factor ANOVA for two-variable confidence.

is because control was important for isolating one variable, but smoothness was not helpful. When looking at relationships between variables, control was still important, but smoothness mattered as well. Another subject observed that slow changes were good for examining detail while fast changes were useful for gaining an overall feel for the data. During the exploratory phase of data analysis, the researcher may ask a wide range of questions about the data. A tool providing multiple representations and a variety of manipulations can help the researcher gain insight into the answers to these questions.

8. Conclusions

Calico, a dynamic tool for the creation and manipulation of color mappings for the exploration of multivariate, quantitative data, was used to study the effects of user control and smooth change on user preference, accuracy, and confidence. The results of the study, as well as other user experiences with Calico, support the hypothesis that dynamic manipulation of color mappings is a useful feature of systems for the exploration of quantitative data using color. The main effect observed so far is a clear user preference for representations providing control over the mapping, a small but significant increase in accuracy, and greater confidence in information gleaned from manipulable displays. A smaller and less consistent effect showed greater user preference for and confidence in representations which provided smooth change between images.

More experimental study is needed to more comprehensively evaluate the effectiveness of data display techniques. Two important directions are the study of actual researchers in more natural data exploration situations and the consideration of more qualitative aspects of data analysis. While it is easier to judge the accuracy of quantitative questions like those asked in this experiment, insight into the qualitative nature of the patterns and interrelationships within the data is probably a more important product of data exploration.

Acknowledgements

I am grateful for the valuable contributions of many people to this research. Brice Tebbs supplied the initial inspiration for Calico and some key parts of its implementation. The software efforts of the whole Pixel-Planes team made the implementation of Calico possible. Greg Turk, Frederick P. Brooks, Jr., and various anonymous reviewers provided insightful suggestions about both this research and this paper. Terry Yoo contributed constant encouragement and truly heroic photography.

This work has been supported in part by grants from the NIH Division of Research Resources (#5-R24-RR-02170) and the Office of Naval Research (#N00014-86-K-0680).

Appendix : Sample Worksheet

The images that you will see show two socioeconomic variables : median age and percentage of workers who drive to work. As in the tutorial, these variables are represented by levels of green and purple. Specifically,

Green = median age
Purple = percentage of workers who drive to work

Each image shows a different balance between the two variables. Each will be shown for a few seconds. When all images have been shown, the display will go back to the beginning of the series.

Please answer the following questions :

1. What is the median age in Dane County, Wisconsin? (Dane County is outlined in black)

How confident are you about this figure?
Not 1 2 3 4 5 6 7 8 9 10 Very
confident confident

2. What percentage of workers drive to work in places where the median age is greater than 45?

How confident are you about this figure?
Not 1 2 3 4 5 6 7 8 9 10 Very
confident confident

3. How correlated do the variables appear to you?

Negatively Not Positively
correlated correlated correlated
5 4 3 2 1 0 1 2 3 4 5

How confident are you about this judgement?
Not 1 2 3 4 5 6 7 8 9 10 Very
confident confident

4. Point out a place that seems interesting to you. Why does it seem interesting?

References

Cox, Donna (1988), Using the Supercomputer to Visualize Higher Dimensions: An Artist's Contribution to Scientific Visualization, *Leonardo : Journal of Art, Science, and Technology*, vol. 22, no. 3, pp. 133-242.

Fuchs, Henry, Jack Goldfeather, Jeff P. Hultquist, Susan Spach, John D. Austin, Frederick P. Brooks, Jr., John G. Eyles, and John Poulton (1985), Fast Spheres, Shadows, Textures, Transparencies, and Image Enhancements in Pixel-Planes, *Computer Graphics*, vol. 19, no. 3, pp. 111-120.

Guitard, Richard and Colin Ware (1990), A Color Sequence Editor, ACM Transactions on Graphics, vol. 9, no. 3, pp. 338-341.

Levkowitz, Haim [1988], Color in Computer Graphic Representation of Two-Dimensional Parameter Distributions, Ph.D. dissertation, University of Pennsylvania.

Pham, Binh (1990), Spline-based Color Sequences for Univariate, Bivariate, and Trivariate Mapping, *Proceedings: Visualization '90*, pp. 202-208.

Pizer, Stephen M., and John B. Zimmerman (1983), Color Display in Ultrasonography, *Ultrasound in Medicine and Biology*, vol. 9, no. 4, pp. 331-345.

Rheingans, Penny, and Brice Tebbs (1990), A Tool for Dynamic Explorations of Color Mappings, *Computer Graphics*, vol. 24, no. 2, pp. 145-146.

Robertson, Philip K. (1988), Visualizing Color Gamuts : A User Interface for the Effective Use of Perceptual Color Spaces in Data Displays, *IEEE Computer Graphics and Applications*, Sept. 1988, pp. 50-64.

Robertson, Philip K., and John F. O'Callaghan (1986), The Generation of Color Sequences for Univariate and Bivariate Mapping, *IEEE Computer Graphics and Applications*, Feb. 1986, pp. 24-32.

Smith, Alvy Ray (1978), Color Gamut Transform Pairs, *Computer Graphics*, vol. 12, no. 3, pp. 12-18.

Trumbo, Bruce E. (1981), A Theory for Coloring Bivariate Statistical Maps, *The American Statistician*, vol. 35, no. 4, pp. 220-226.

Wainer, Howard, and Carl M. Francolini (1980), An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps, *The American Statistician*, vol. 34, no. 2, pp. 81-93.

Ware, Colin (1988), Color Sequences for Univariate Maps : Theory, Experiments and Principles, *IEEE Computer Graphics and Applications*, Sept. 1988, pp. 41-49.

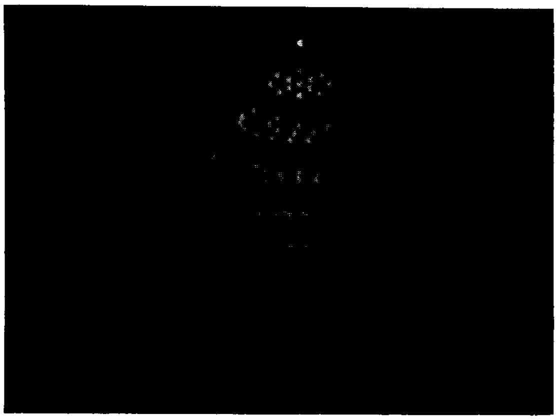

Figure 1: Color path.

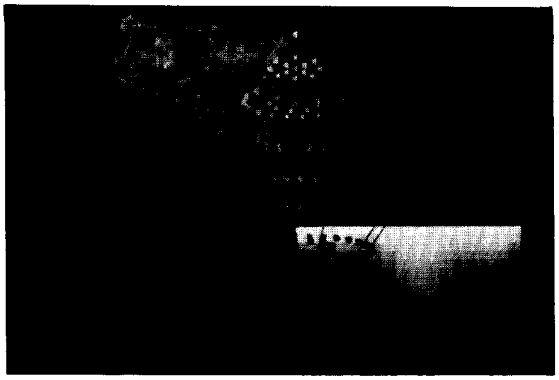

Figure 2: Calico display. Median family income is displayed using hue, while average education level is displayed using brightness.

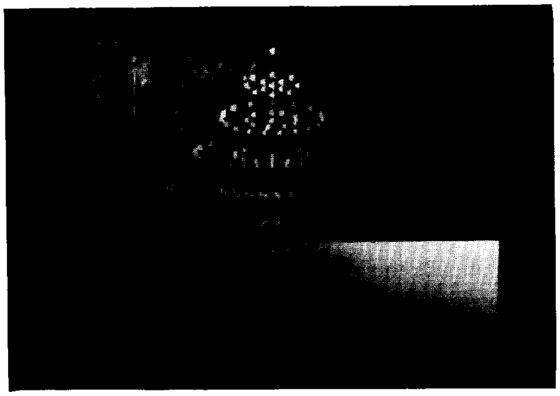

Figure 4: Sample stimulus image. Income is displayed by levels of green, while education is displayed by levels of purple. The mapping is described by the equations

Hue = 0.35
Lightness = u*S + v*(1-S)
Saturation = 2*(u*S - v*(1-S))

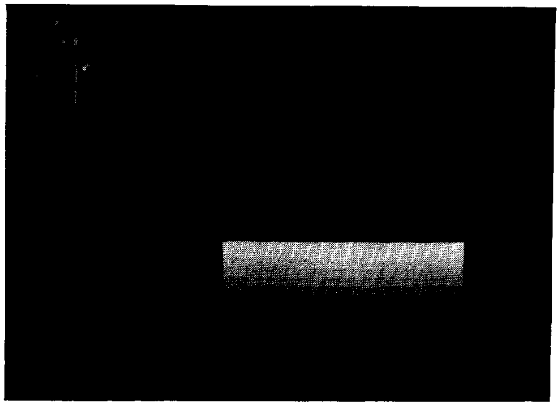

Figure 5a.

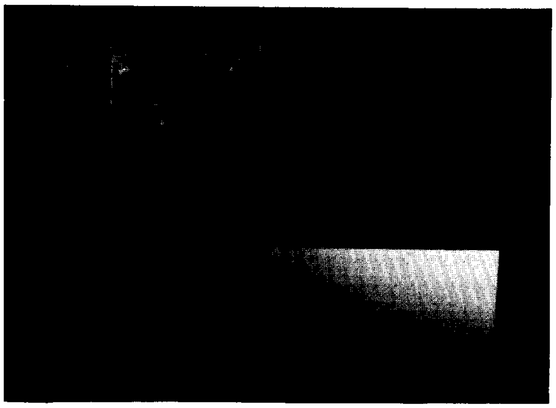

Figure 5b.

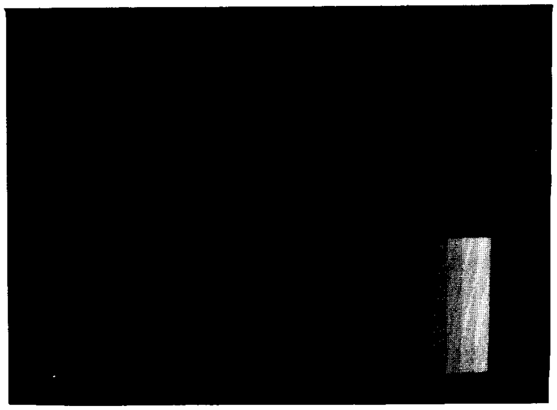

Figure 5c.

Figure 5 (a-c): Four levels of relative variable contribution:
a) shows just the contribution of education level,
b) balanced contributions of education and income,
c) just the contribution of income level.

(See color plates, p. CP-28.)