Data Visualization Principles: Color

CSC444

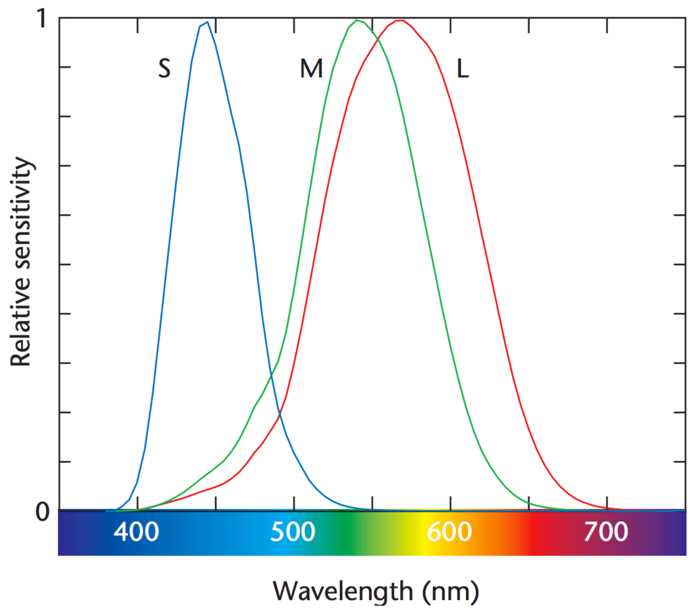
Acknowledgments for today's lecture:
Tamara Munzner, Miriah Meyer, Maureen Stone

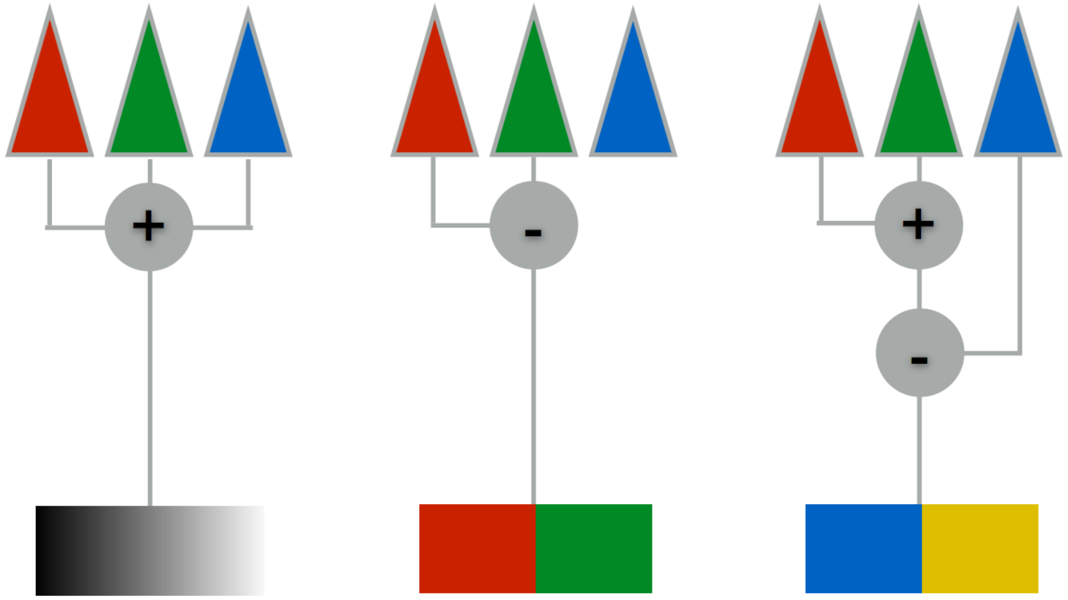
ANNOUNCEMENTS

Assignment 3 solution
Assignment 4 due tonight
Assignment 5 posted

https://cscheid.net/projects/d3-drills/

RECAP



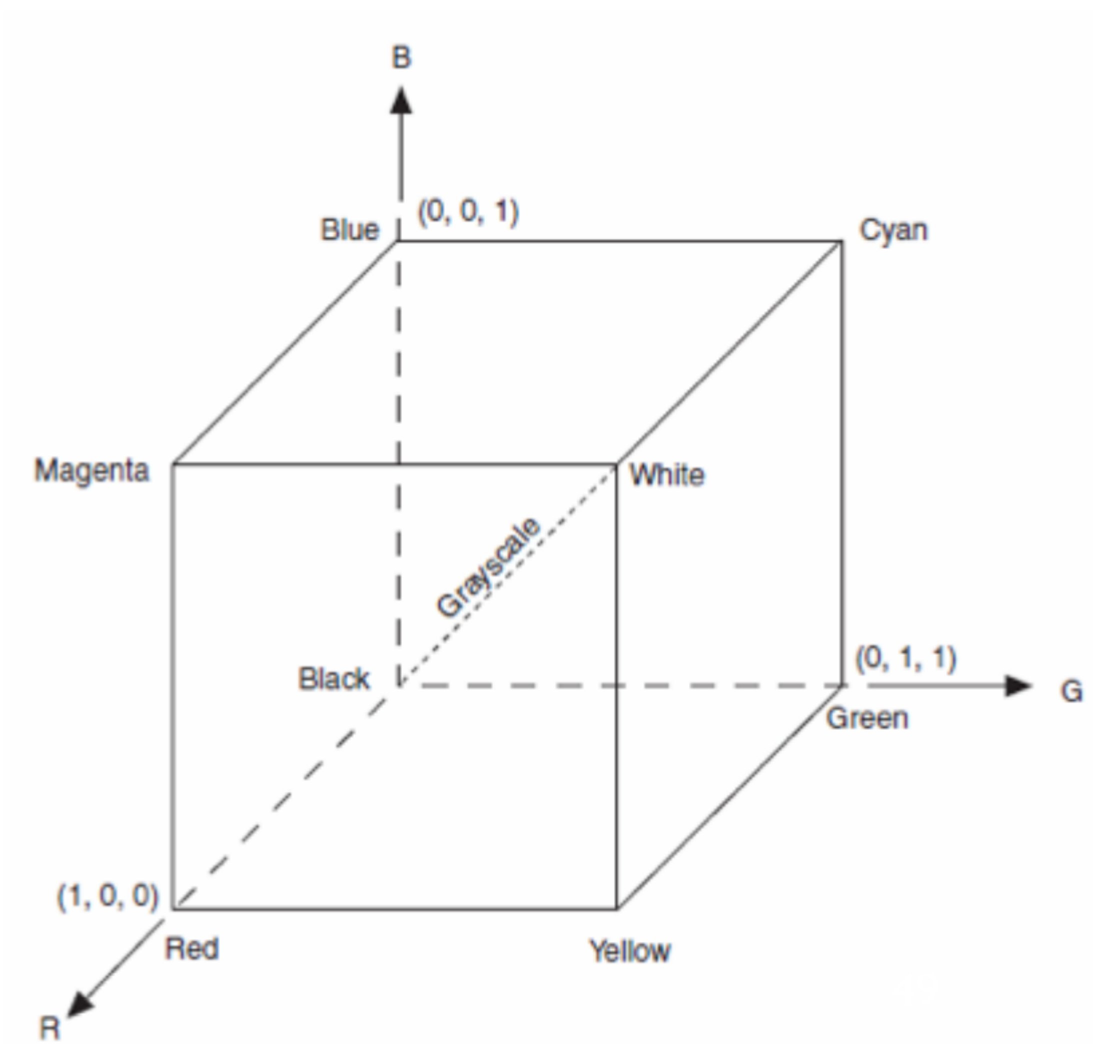


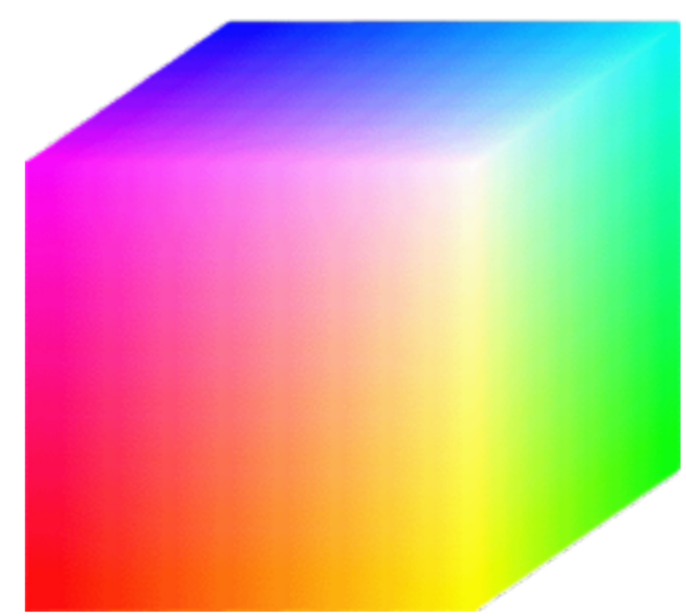
COLOR SPACES

DEVICE DEPENDENT

RGB

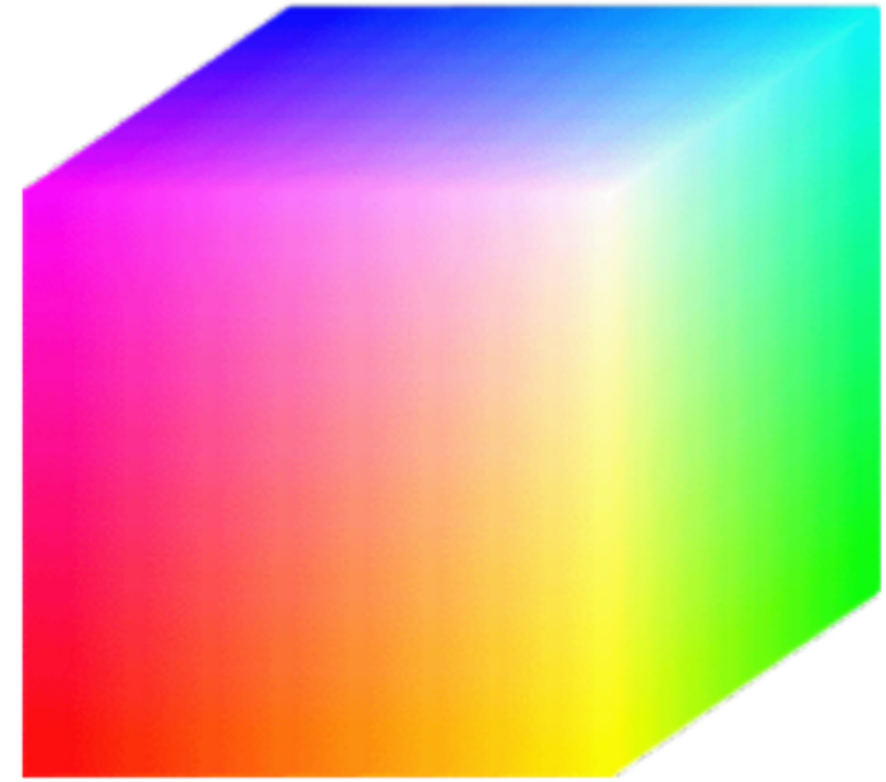
- Device-centric
- What programs want,
  not what humans want

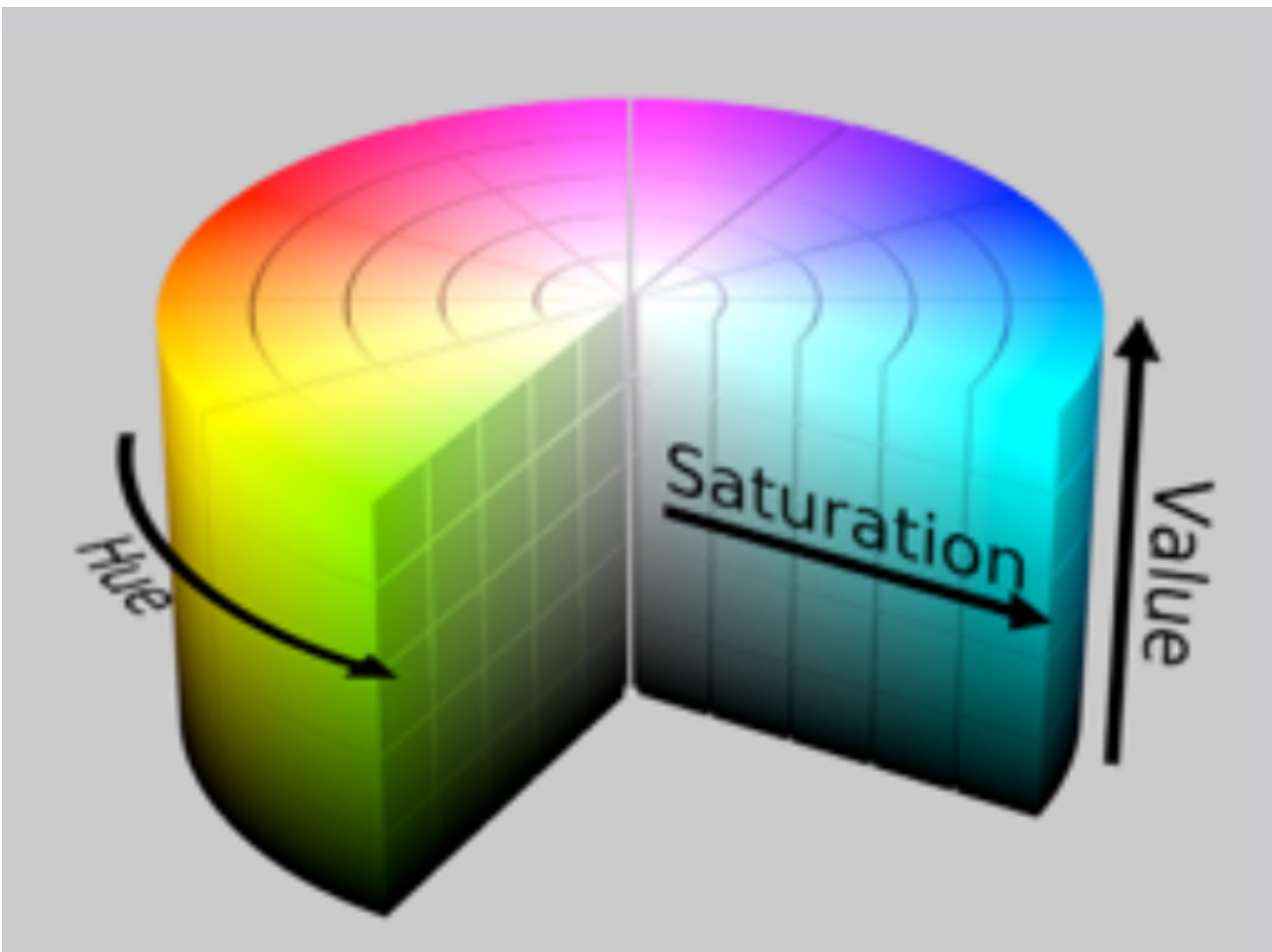




HSV

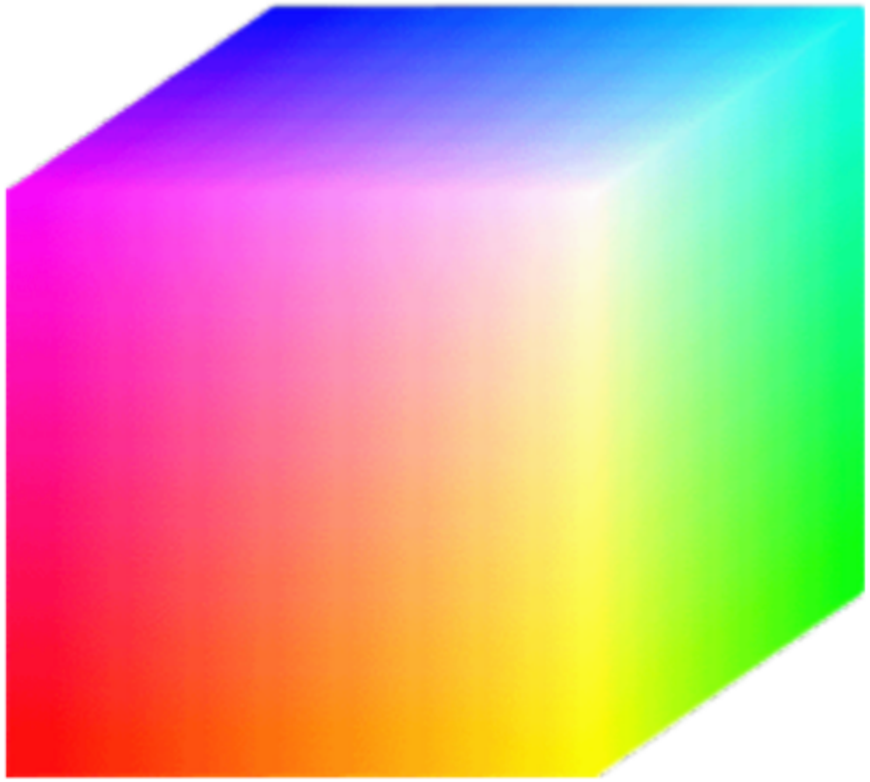
- Still device-centric

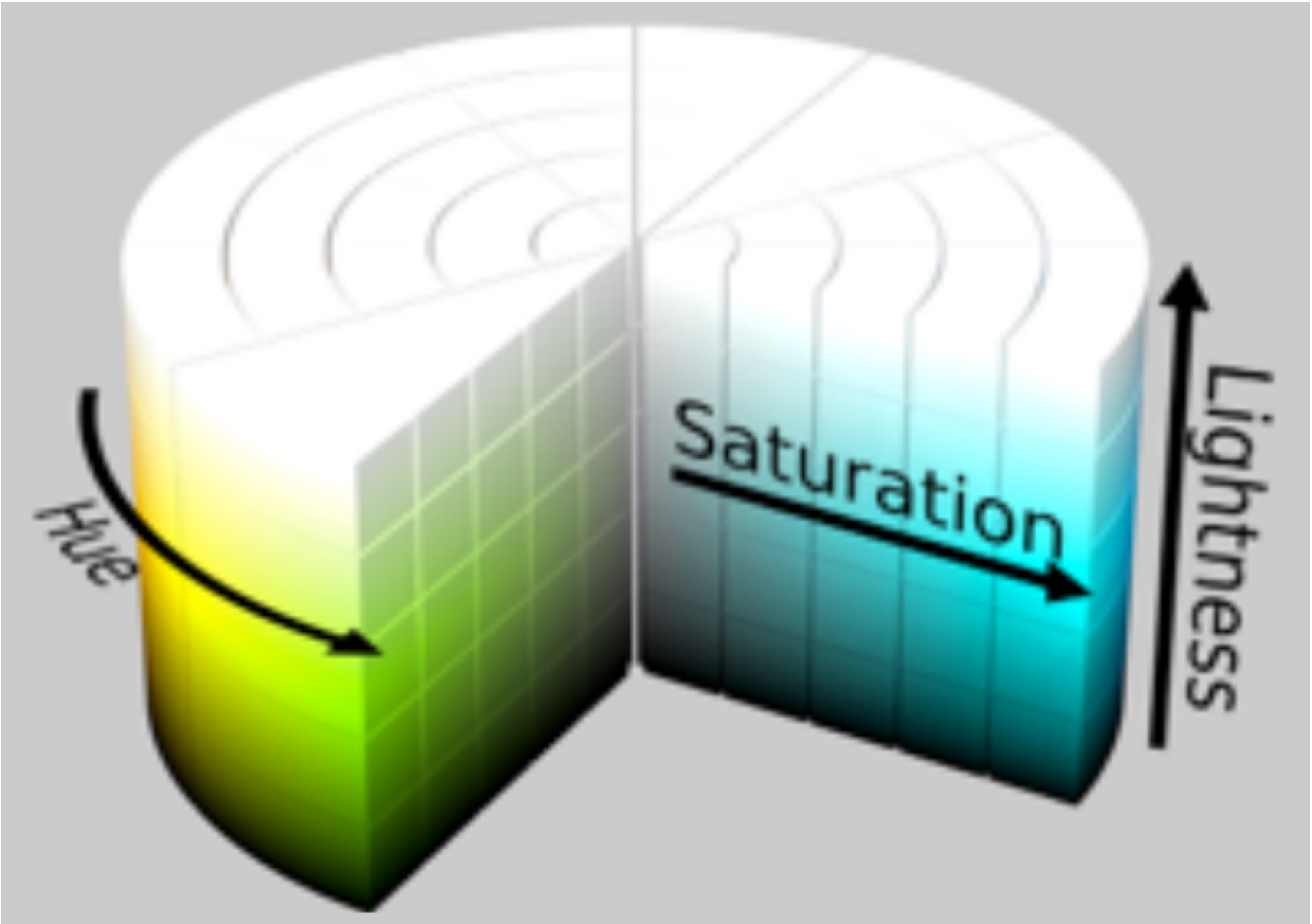




HSL

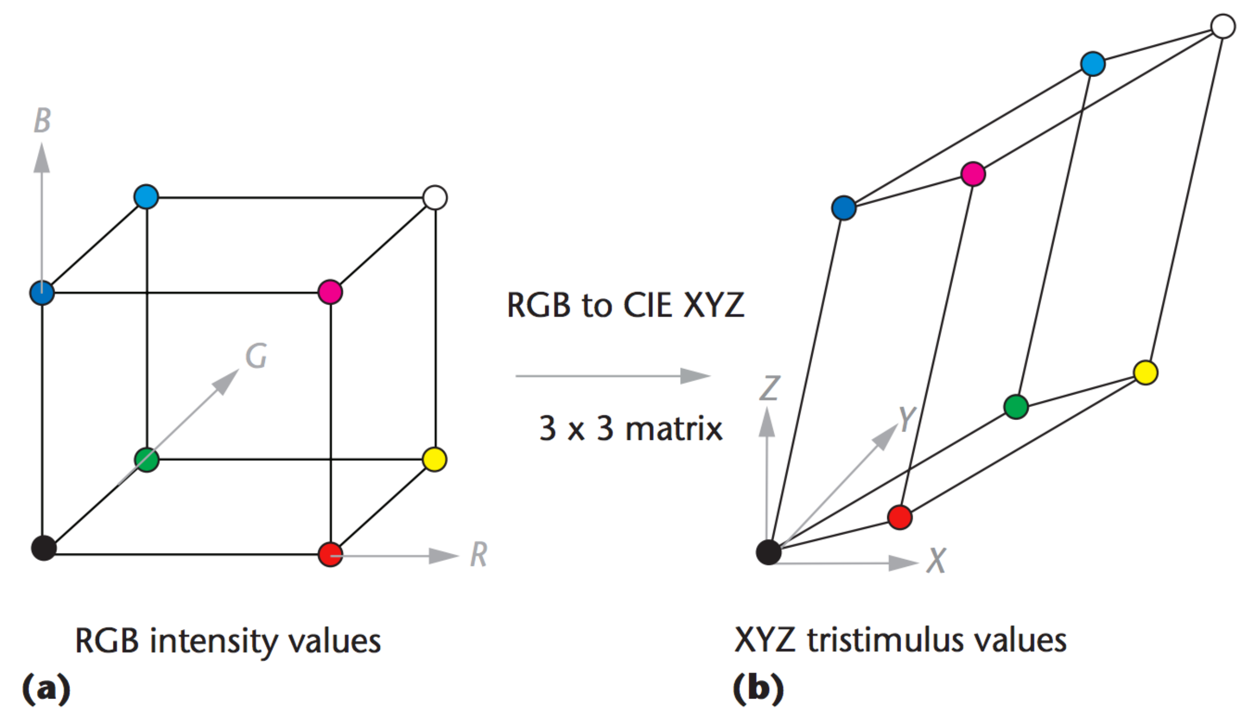
- Still device-centric
- (supported in d3)





DEVICE INDEPENDENT

XYZ Color Space



- "Optically linear"
- CIE designed three **reference spectra: X, Y, Z**
- Designed so that all visible colors have positive coordinates, and Y is "luminance"

Lab Color Space

- “Perceptually uniform”
    - Euclidean distance corresponds, roughly, to perceptual distance (**very useful!**)

Polar Lab (or HCL)

- “Perceptually uniform”, like Lab
- Transform ab to polar coordinates: radius is Chroma, Angle is Hue
- Conversion to/from RGB is complicated, but distances in HCL make sense, **and** it makes sense for humans
  - **Like HSV, but good. All else being equal, think HCL first**

Demos

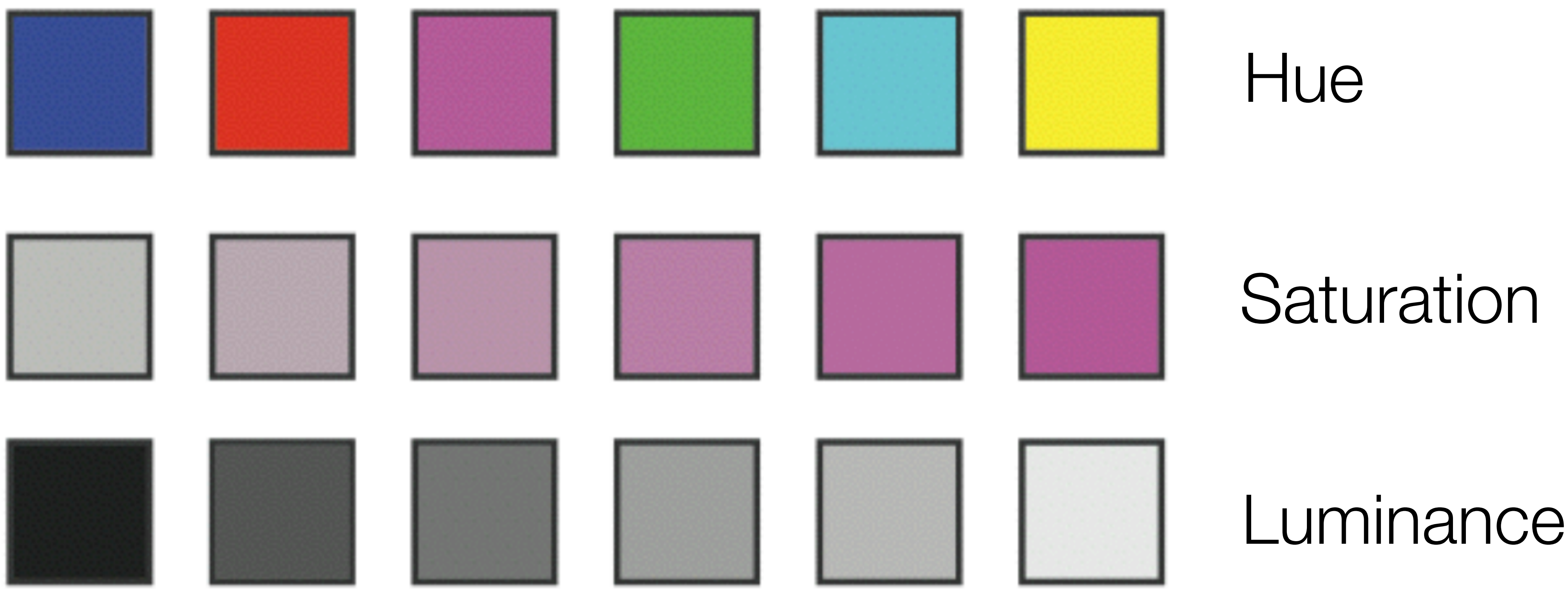
http://cscheid.net/static/20120216/hsv_frame.html

http://cscheid.net/static/20120216/xyz_frame.html

http://cscheid.net/static/20120216/luv_frame.html

http://cscheid.net/static/20120216/hcl_frame.html

Let's use consistent names in class

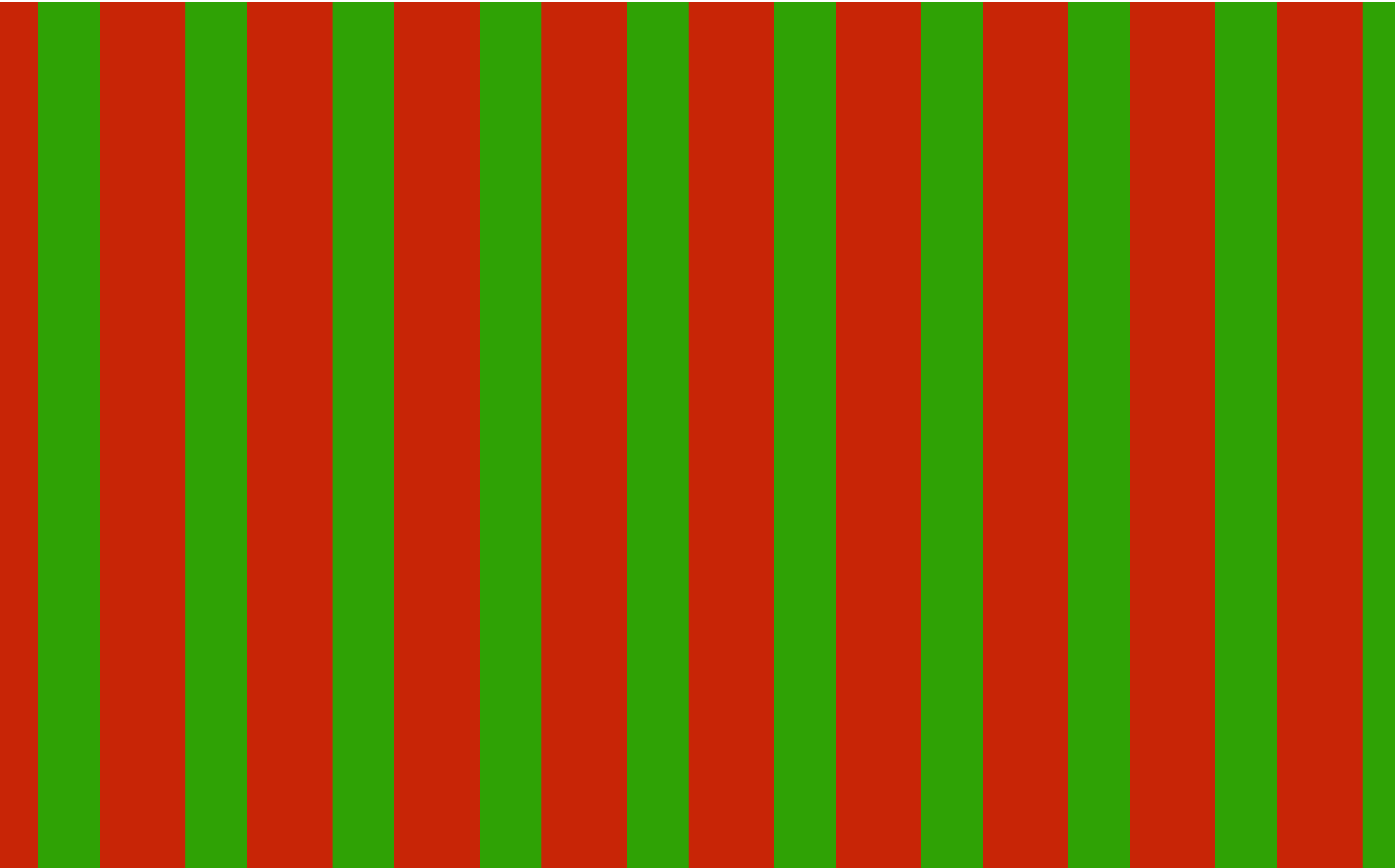


CONSEQUENCES FOR DESIGN

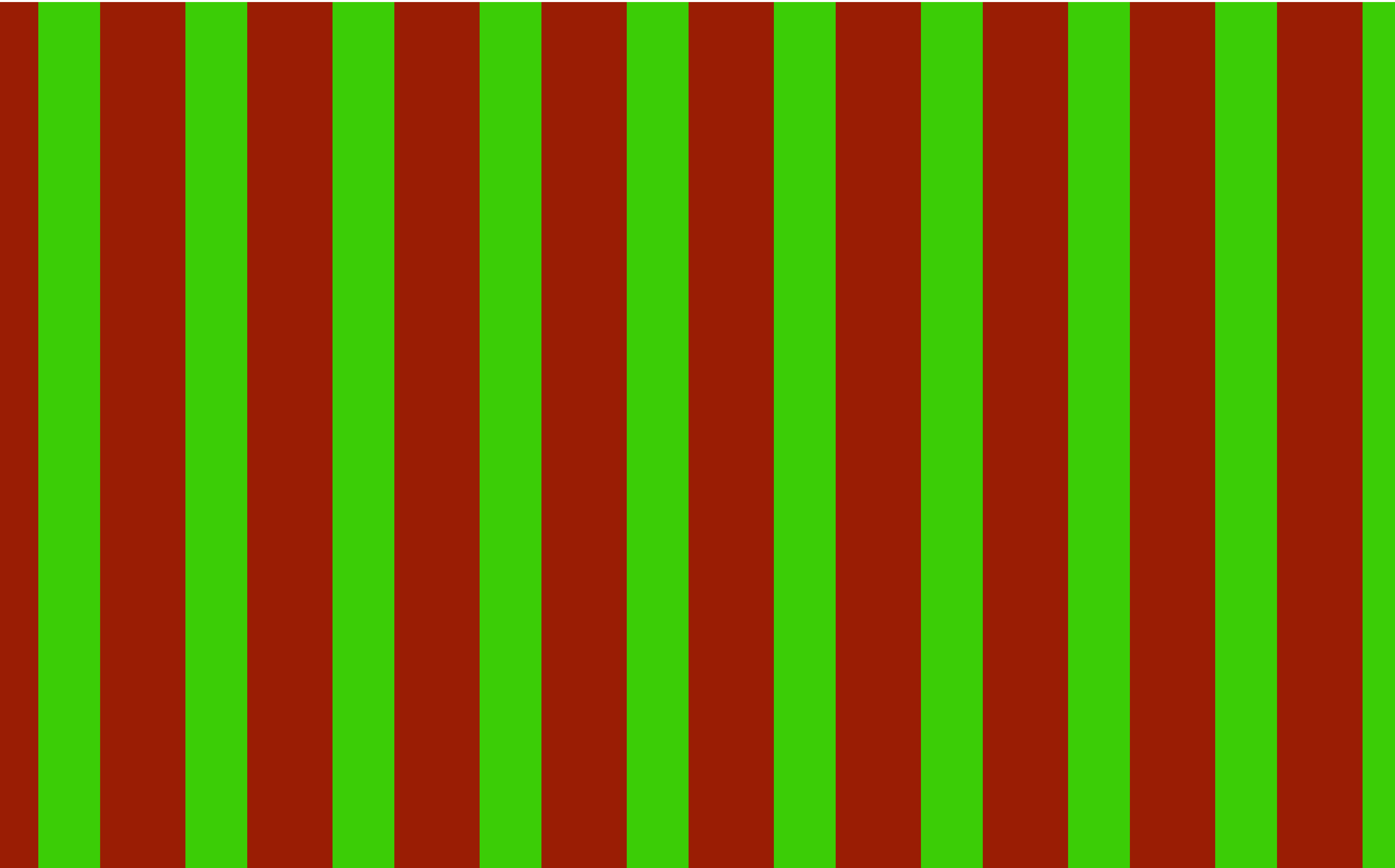
“Get it right in black and white”

–Maureen Stone

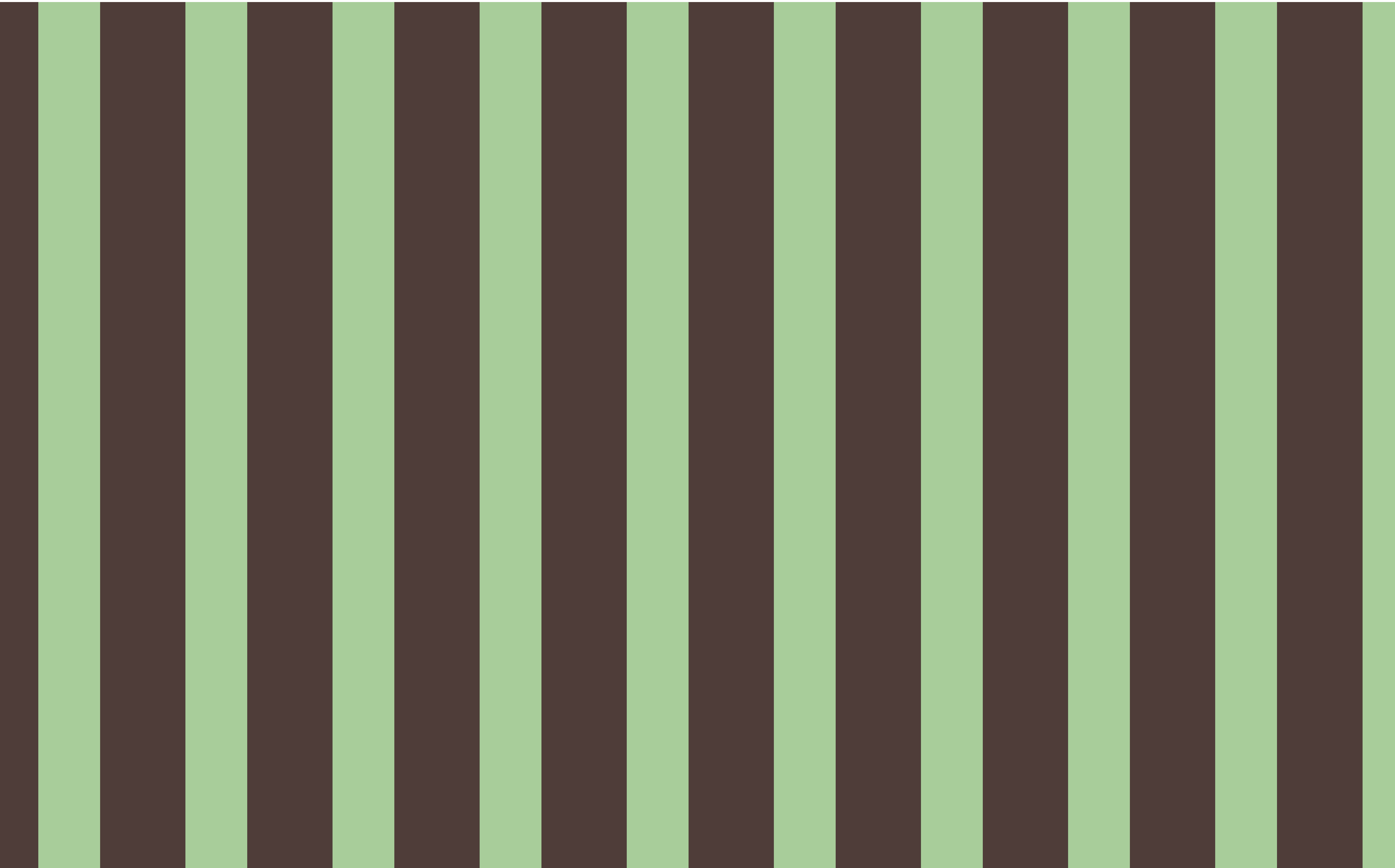
If you're going to show shape variation, do it with luminance

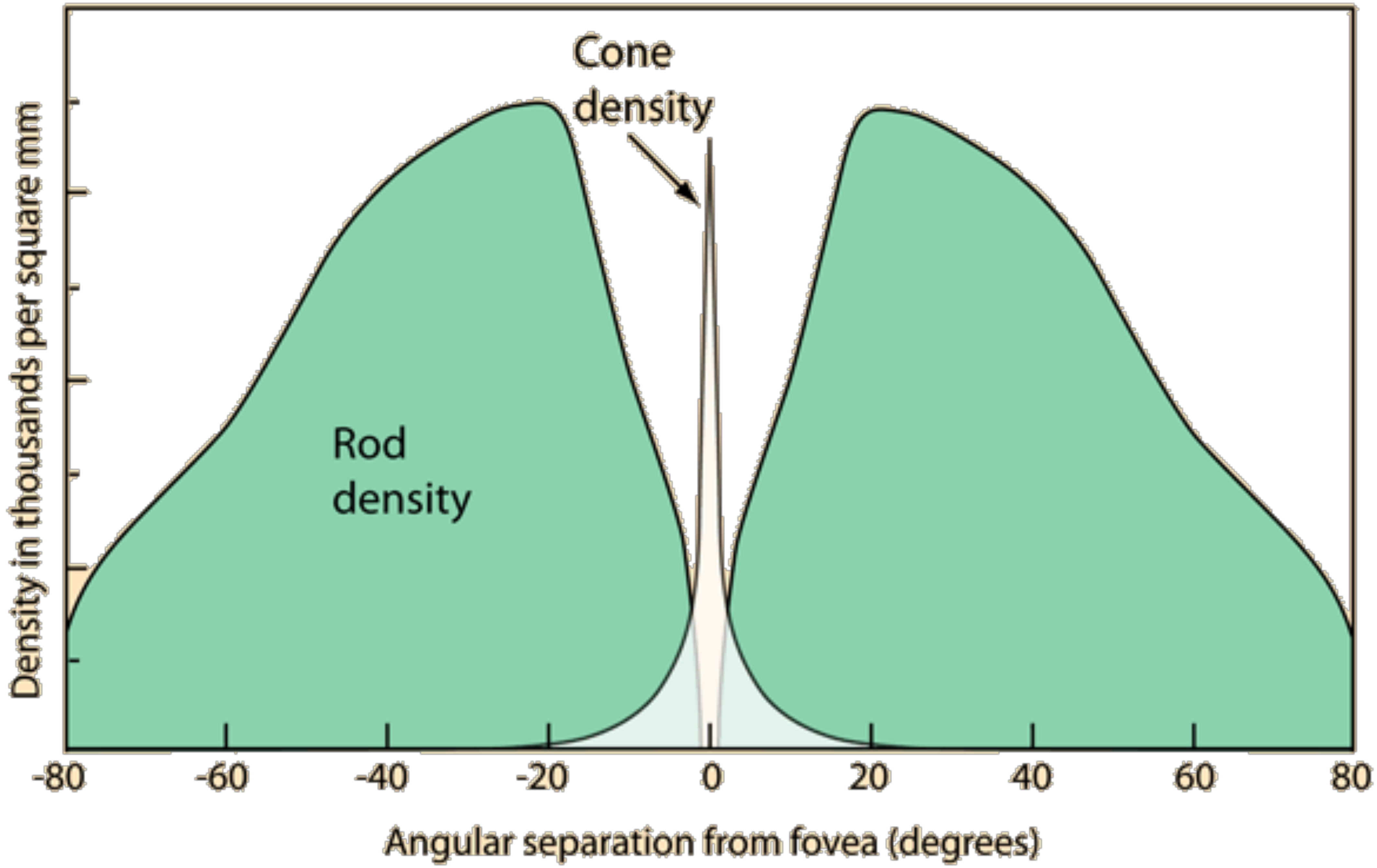


If you're going to show shape variation, do it with luminance

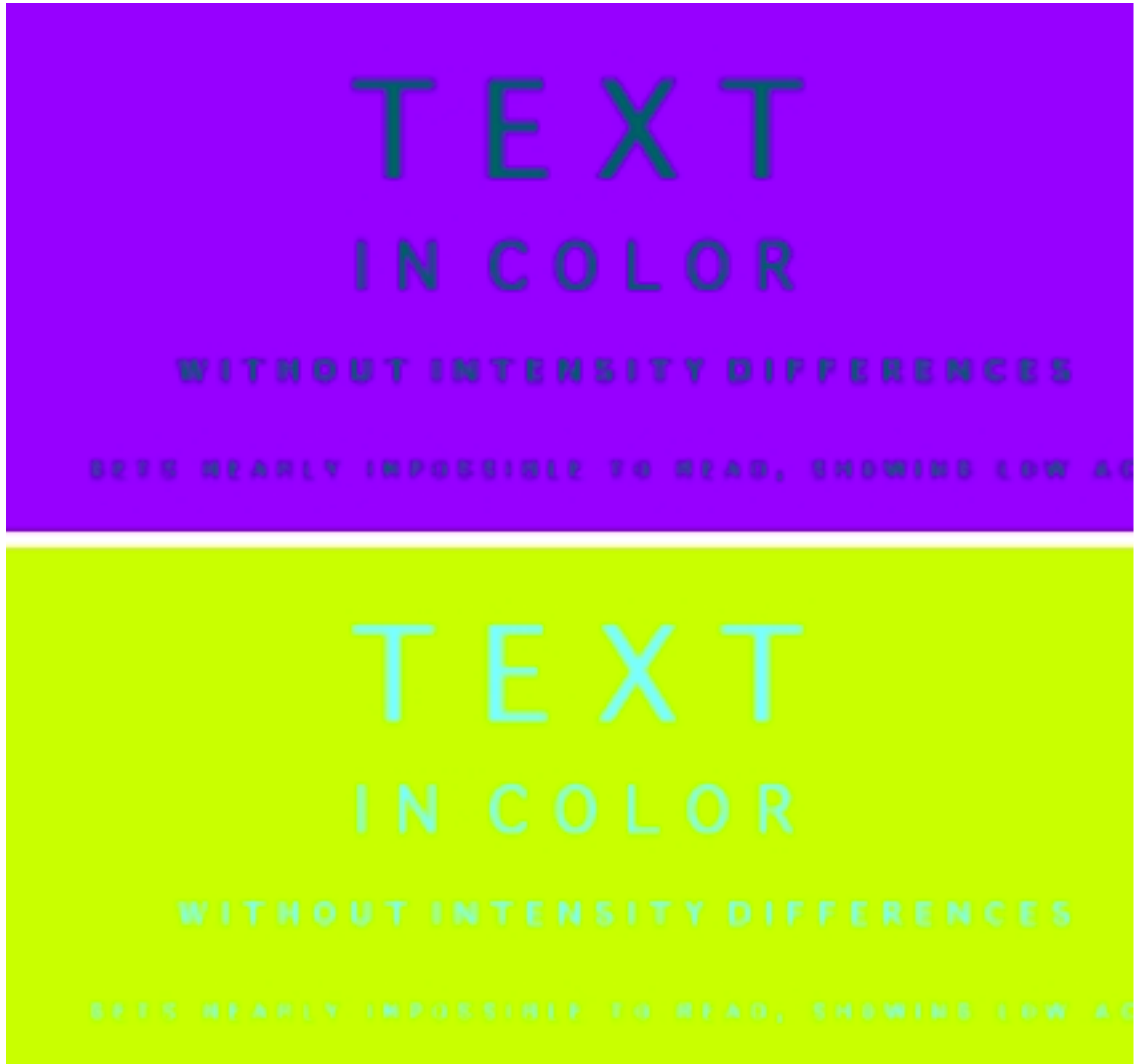


If you’re going to show shape variation, do it with luminance



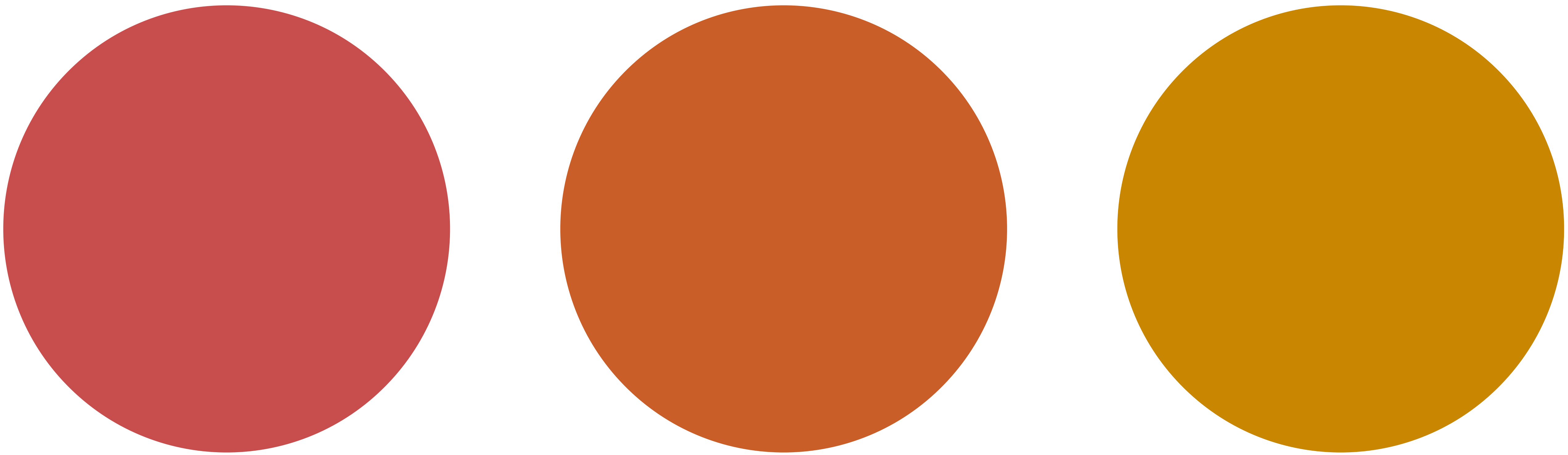


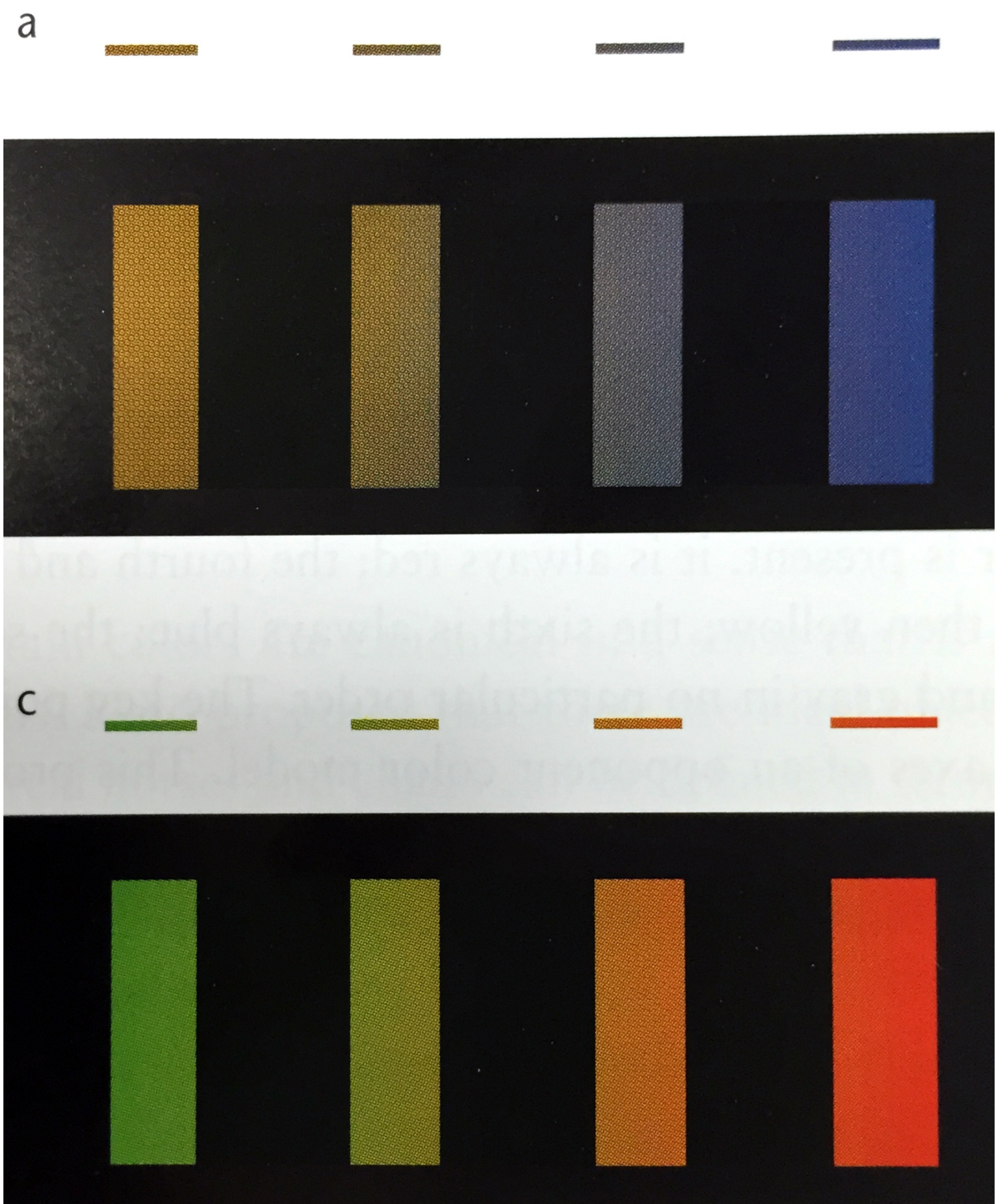
(You can see stars better by looking away from them!)


TEXT
IN COLOR
WITHOUT INTENSITY DIFFERENCES
GETS NEARLY IMPOSSIBLE TO READ, SHOWING LOW AC
TEXT
IN COLOR
WITHOUT INTENSITY DIFFERENCES
GETS NEARLY IMPOSSIBLE TO READ, SHOWING LOW AC

Do not rely only on hue boundaries to depict shape

Do not rely only on hue boundaries to depict shape





Ware, Chapter 4

Area affects saturation perception

A Color-Caused Optical Illusion on a Statistical Graph

WILLIAM S. CLEVELAND and ROBERT McGILL*

Despite the great increase in the use of color in statistical graphics, we know very little about how color affects people's perception of the quantitative information on graphical displays. Perceptual psychologists have already demonstrated that color can cause optical illusions of various kinds. We ran a simple experiment to see if this can happen with a statistical map and found that an illusion did occur.

KEY WORDS: Statistical map; Psychophysics; Bootstrap; Computer graphics; Barycentric plot.

1. INTRODUCTION

Color is being used more and more in statistical graphics. The availability of color output devices for computers has greatly increased, and with easy access to these devices has come an enormous increase in the use of color graphics in the mass media, business reports, and government publications.

Scientific journals are turning to color. In the 10 April 1981 issue of *Science*, for example, substantial use is made of color: a graphical display in which two sets of data are distinguished by plotting one in black and one

areas on each map, and the experimental data were used to determine whether some colors caused areas to look bigger than others. This article describes the experiment and the results of the data analysis.

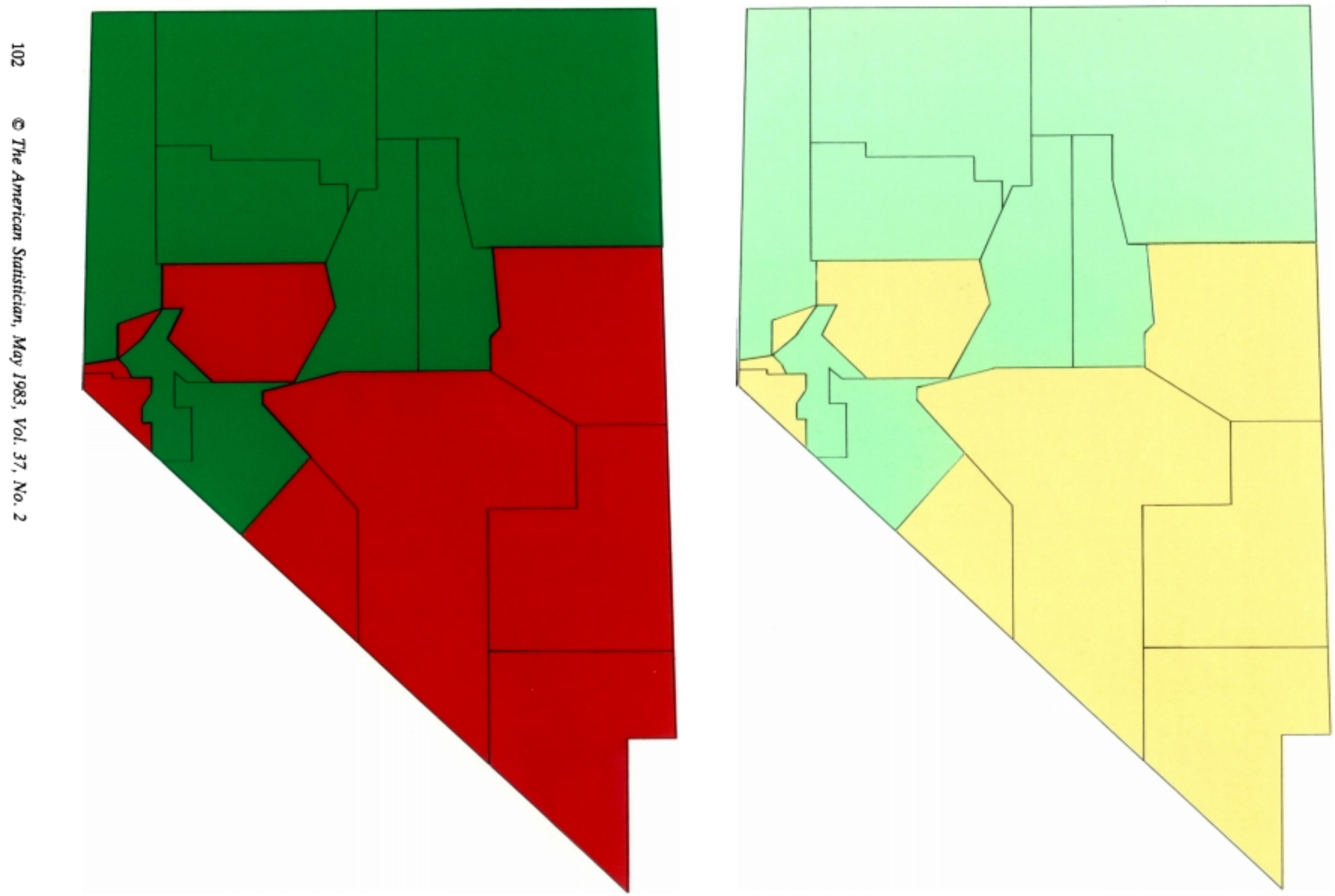
2. EXPERIMENTAL STIMULI AND PROCEDURES

The basic stimulus was a map of Nevada with county boundaries. The geometry of this map is not overly complex—most boundaries are essentially straight lines—yet the sizes and configurations of the counties are not sufficiently regular that estimates of areas could be made by counting them, as might be done on a map of Kansas or Iowa. The number of counties, 17, allows reasonably easy judgment. In no case was a subject required to mentally sum more than 10 areas.

Each stimulus in the experiment was a two-color map with the total area of the counties that were coded by one color very nearly equal to the total area of those that were coded by the second color. One stimulus is shown in Figure 1. There were 10 such divisions of counties into two groups, and in no case did the difference of the areas exceed two square miles.

From the 10 divisions of counties into two groups, four sets of stimuli were generated, each set with the 10 maps previously described. In the first set one group of

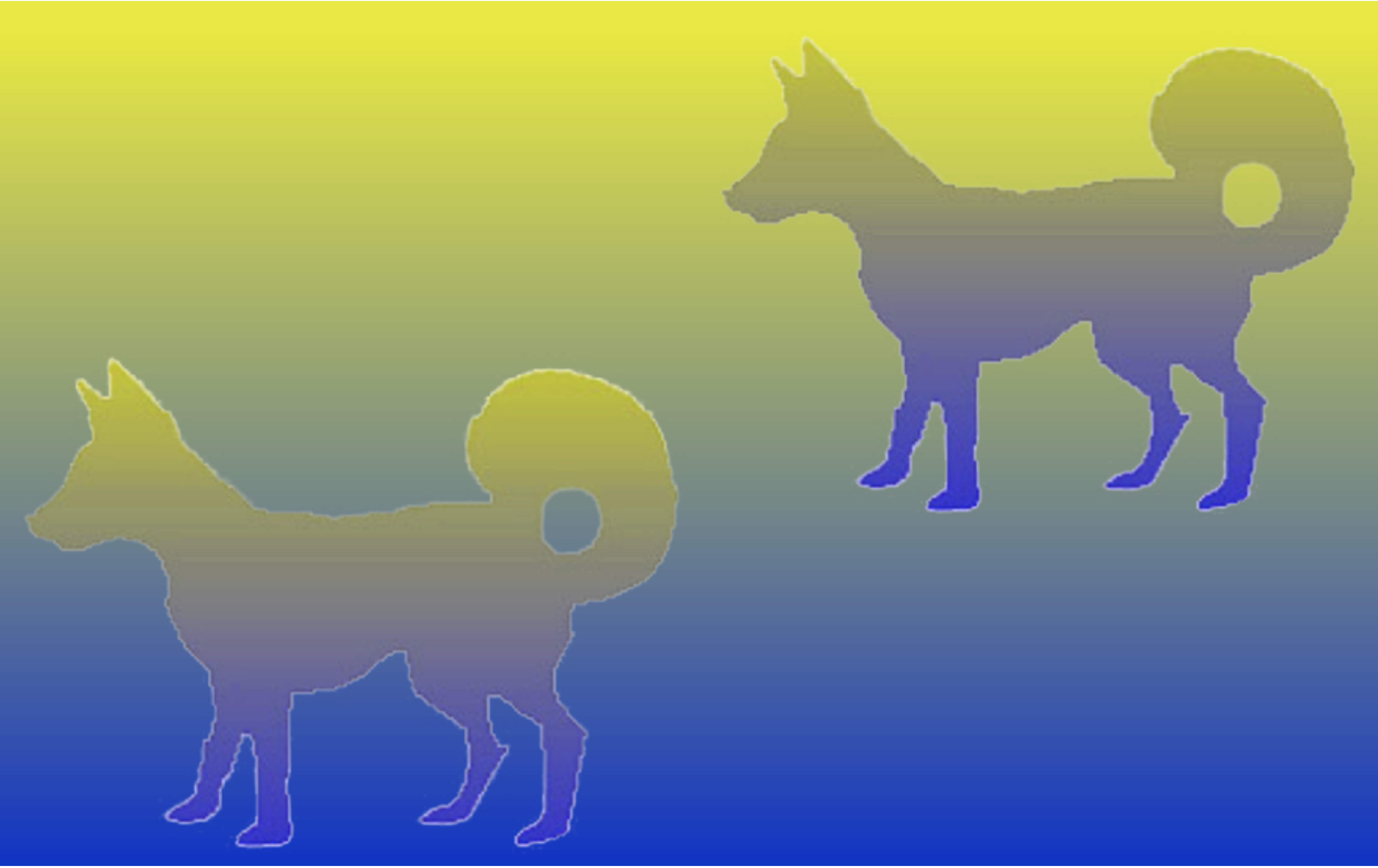
Saturation affects area perception

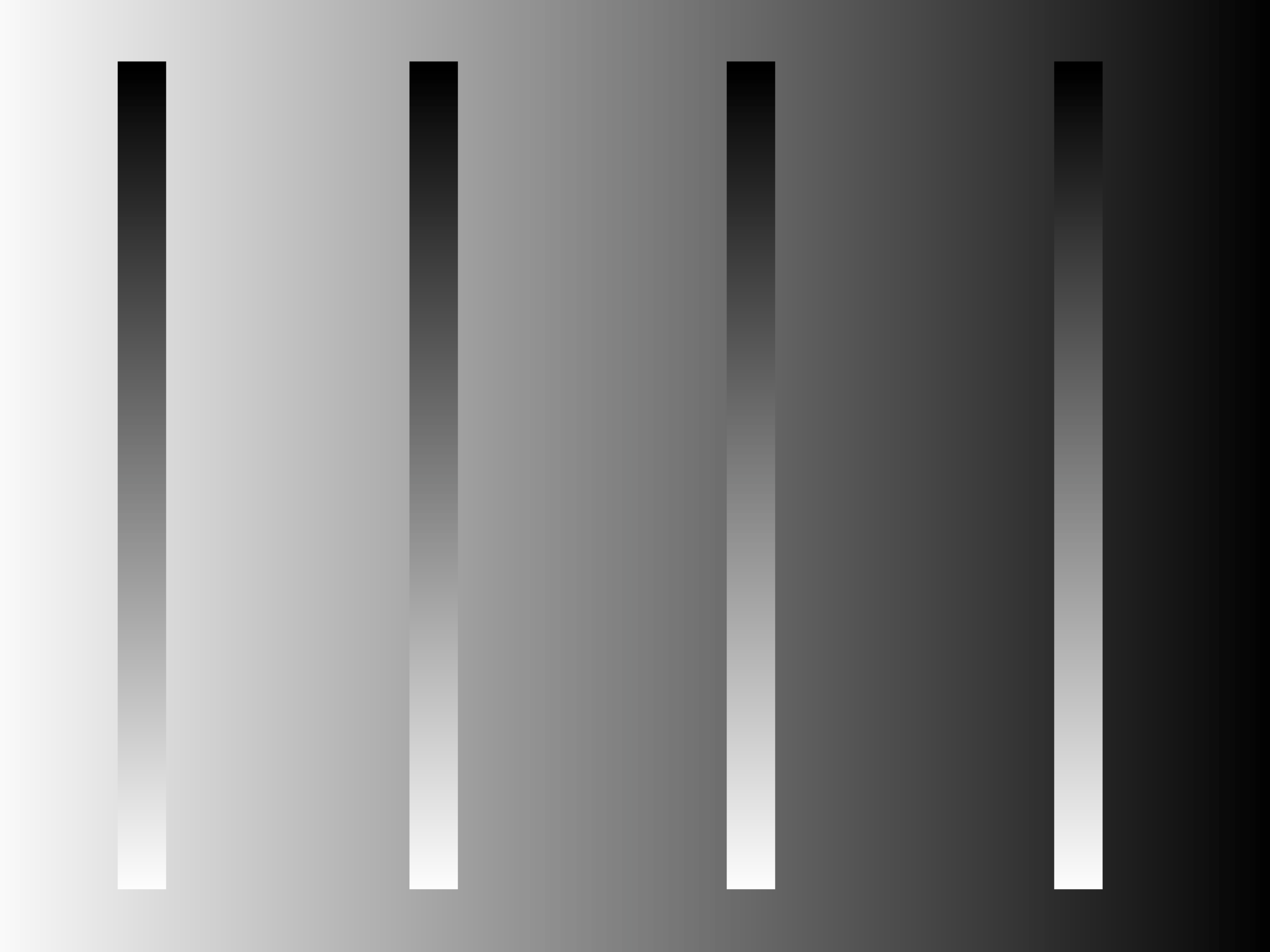


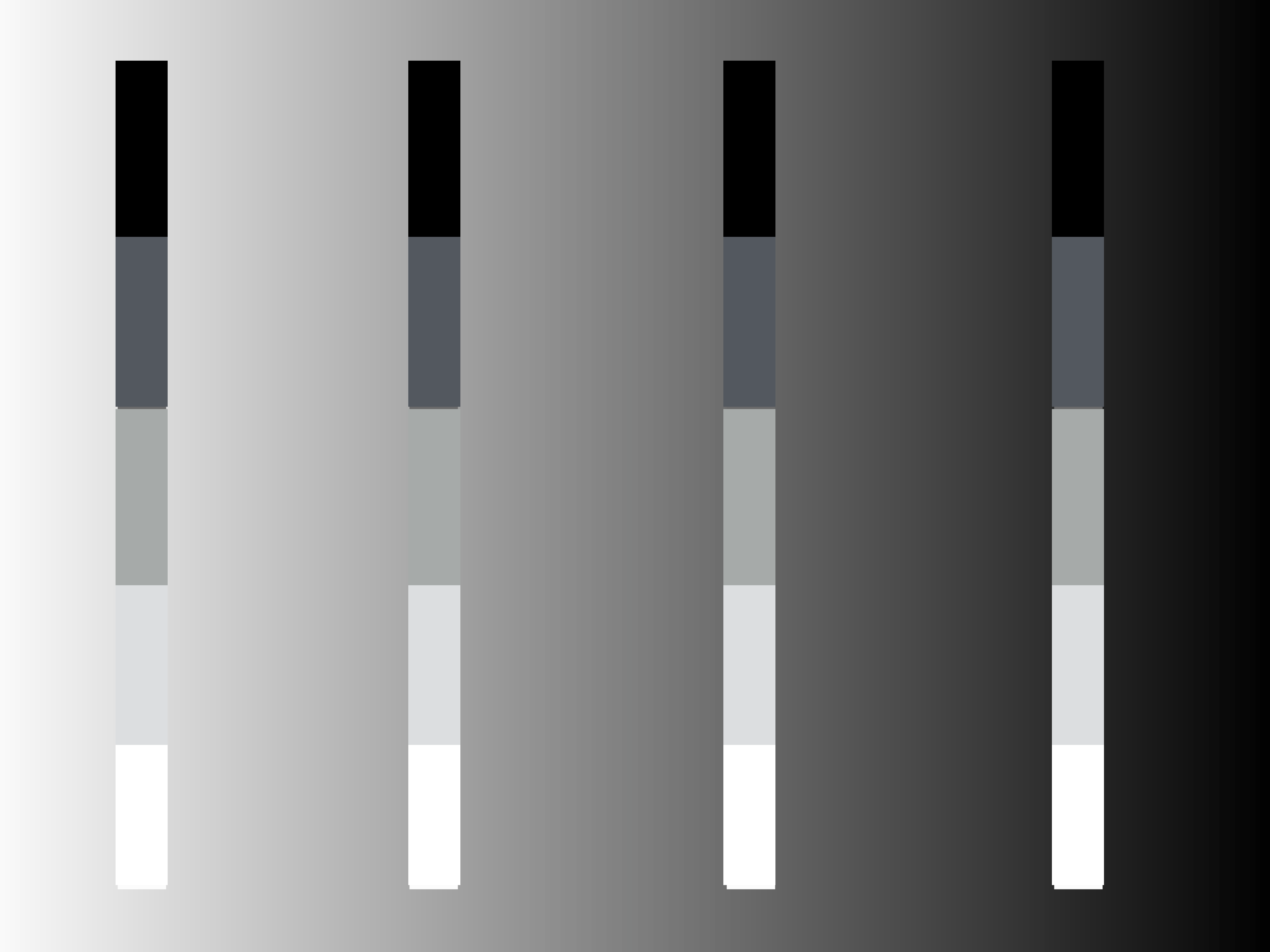
Area affects saturation perception

Saturation affects area perception

Imagine the mess if you try to use both…





Simultaneous contrast is a problem

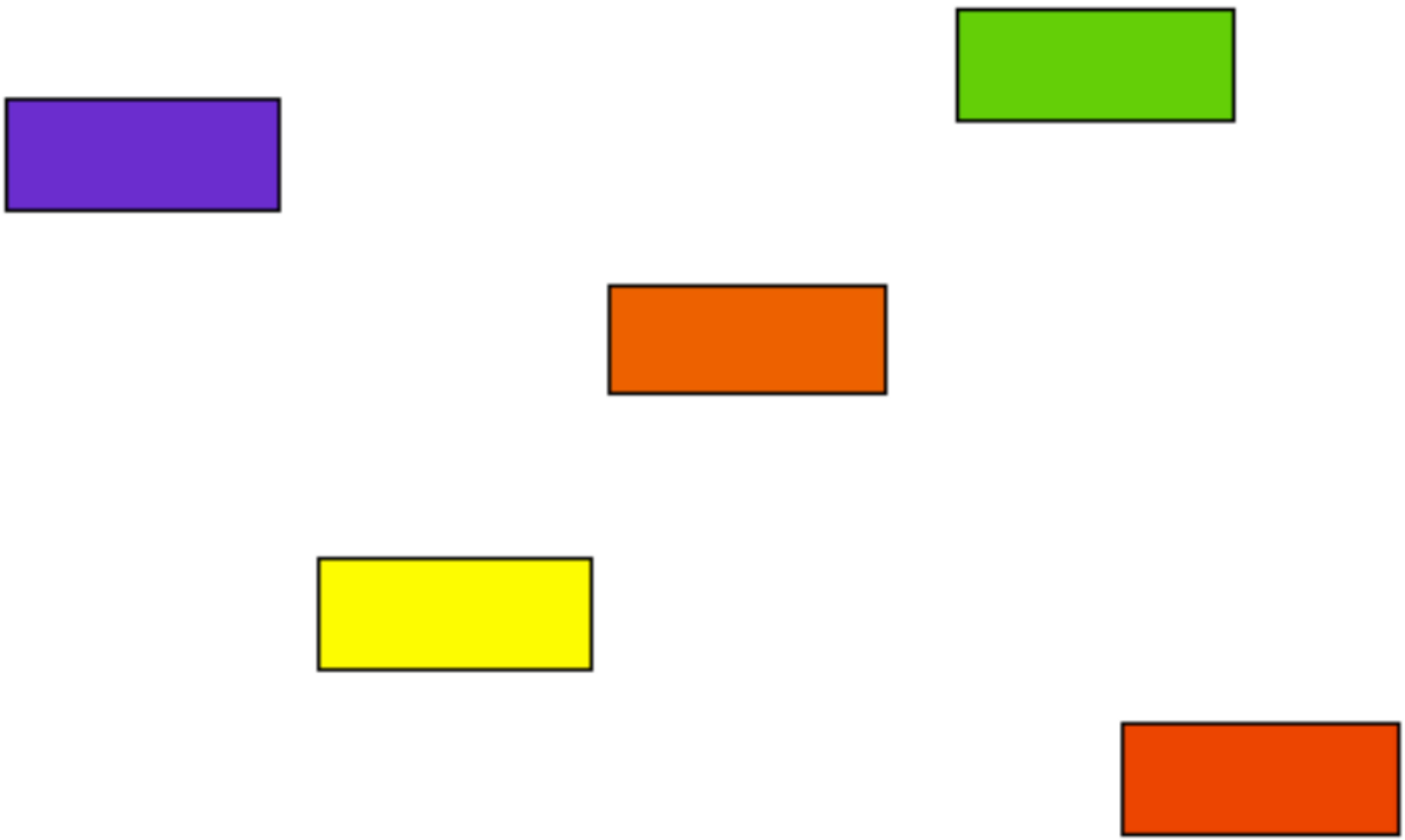
Quantize the plot if background is non-constant
(This comes at a fidelity cost for the data)

“Categorical” data

- Sometimes there’s no implied relationship between different levels of a variable
    - Stimuli must look different, but “**only** different”

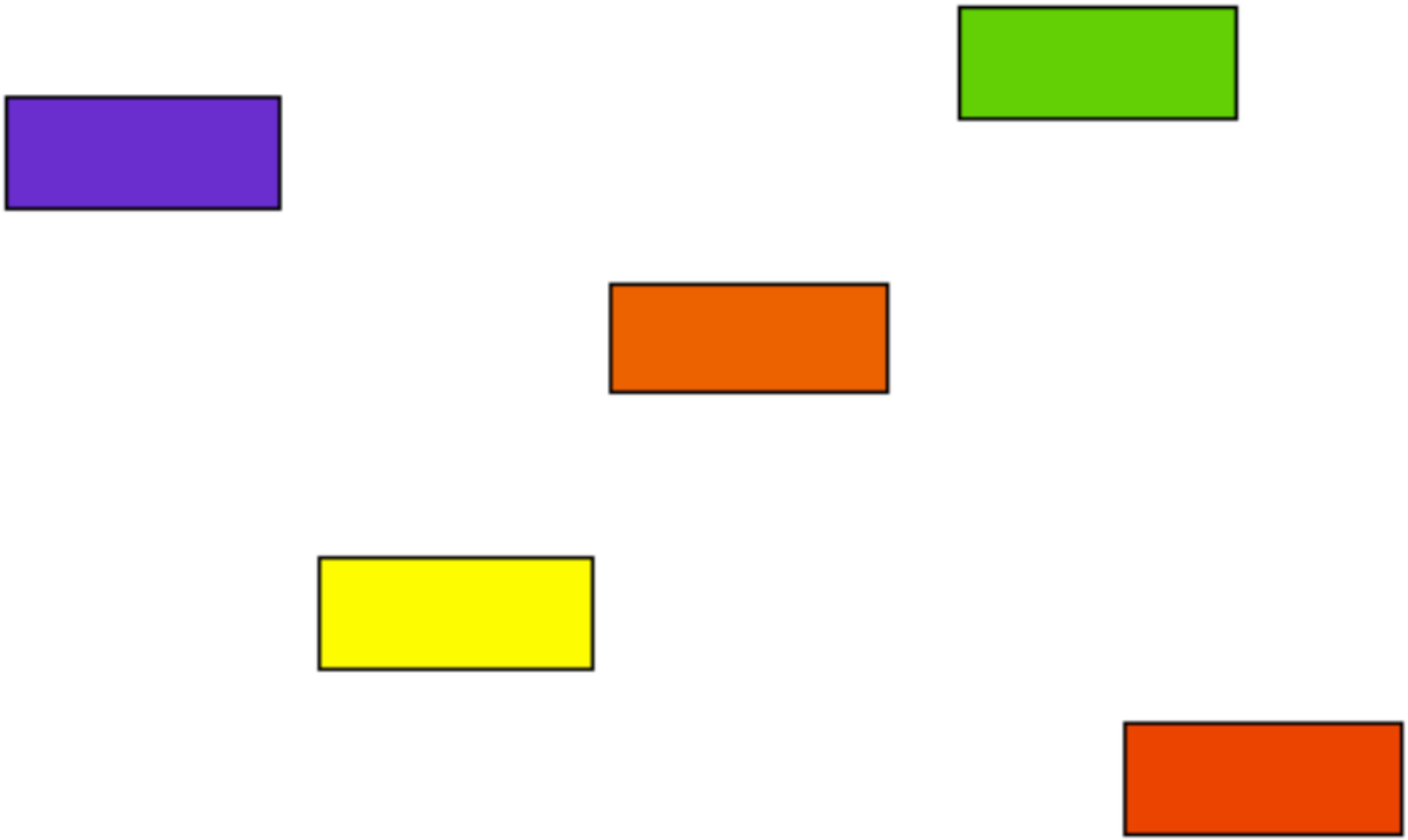
```
d3.scaleOrdinal(d3.schemeCategory10)
```

Order these colors!



You can't…

Order these colors!



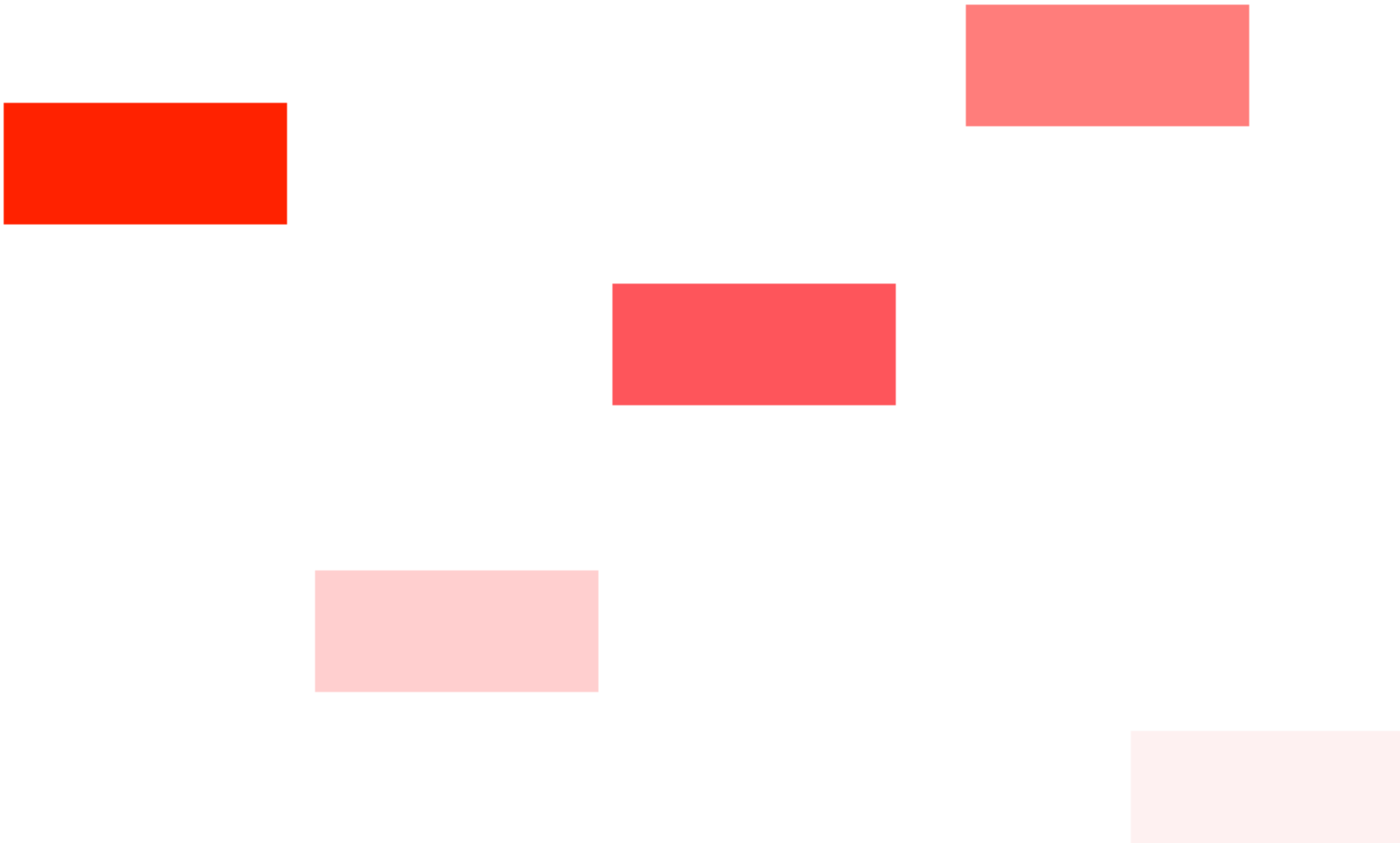
Order these colors!

You can't help but…

Order these colors!

You can't help but…

Order these colors!



Be aware of implied and perceptually forced color relationships

For categorical data, use color only when you have few categories (less than 10)

Editor:
Theresa-Marie Rhyne

Rainbow Color Map (Still) Considered Harmful

David Borland and Russell M. Taylor II
University of North Carolina at Chapel Hill

Research has shown that the rainbow color map is rarely the optimal choice when displaying data with a pseudocolor map. The rainbow color map confuses viewers through its lack of perceptual ordering, obscures data through its uncontrolled luminance variation, and actively misleads interpretation through the introduction of non-data-dependent gradients.

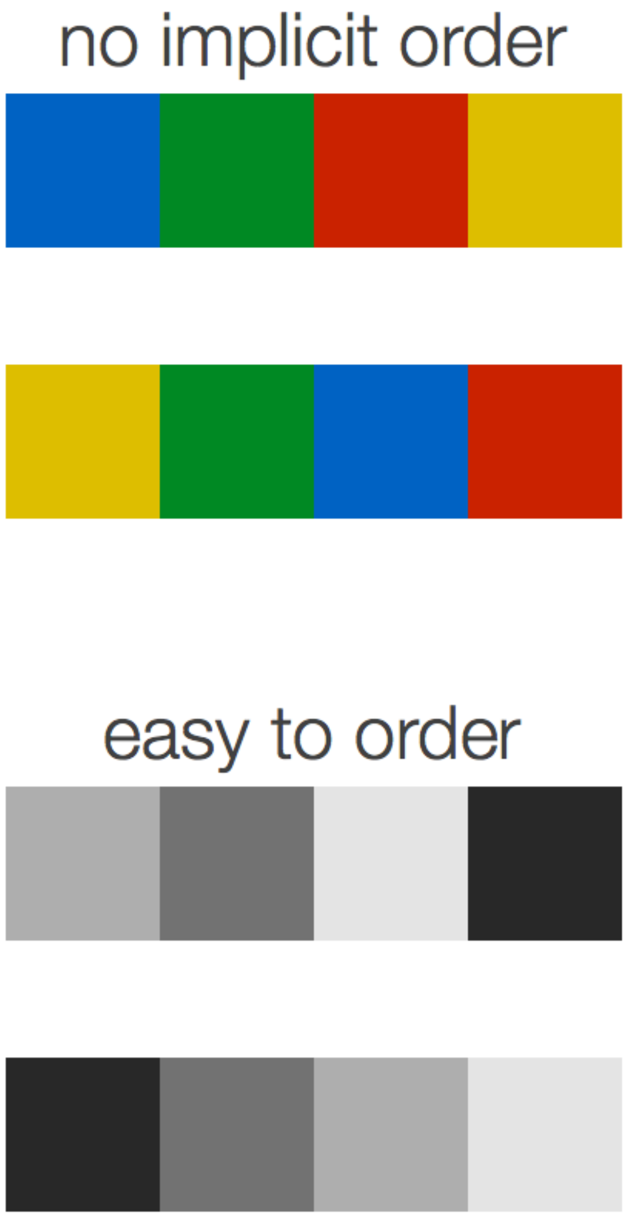
Despite much published research on its deficiencies, the rainbow color map is prevalent in the visualization community. We present survey results showing that the rainbow color map continues to appear in more than half of the relevant papers in IEEE Visualization Conference proceedings; for example, it appeared on 61 pages in 2005. Its use is encouraged by its selection as the default color map used in most visualization toolkits that we inspected. The visualization community must do better.

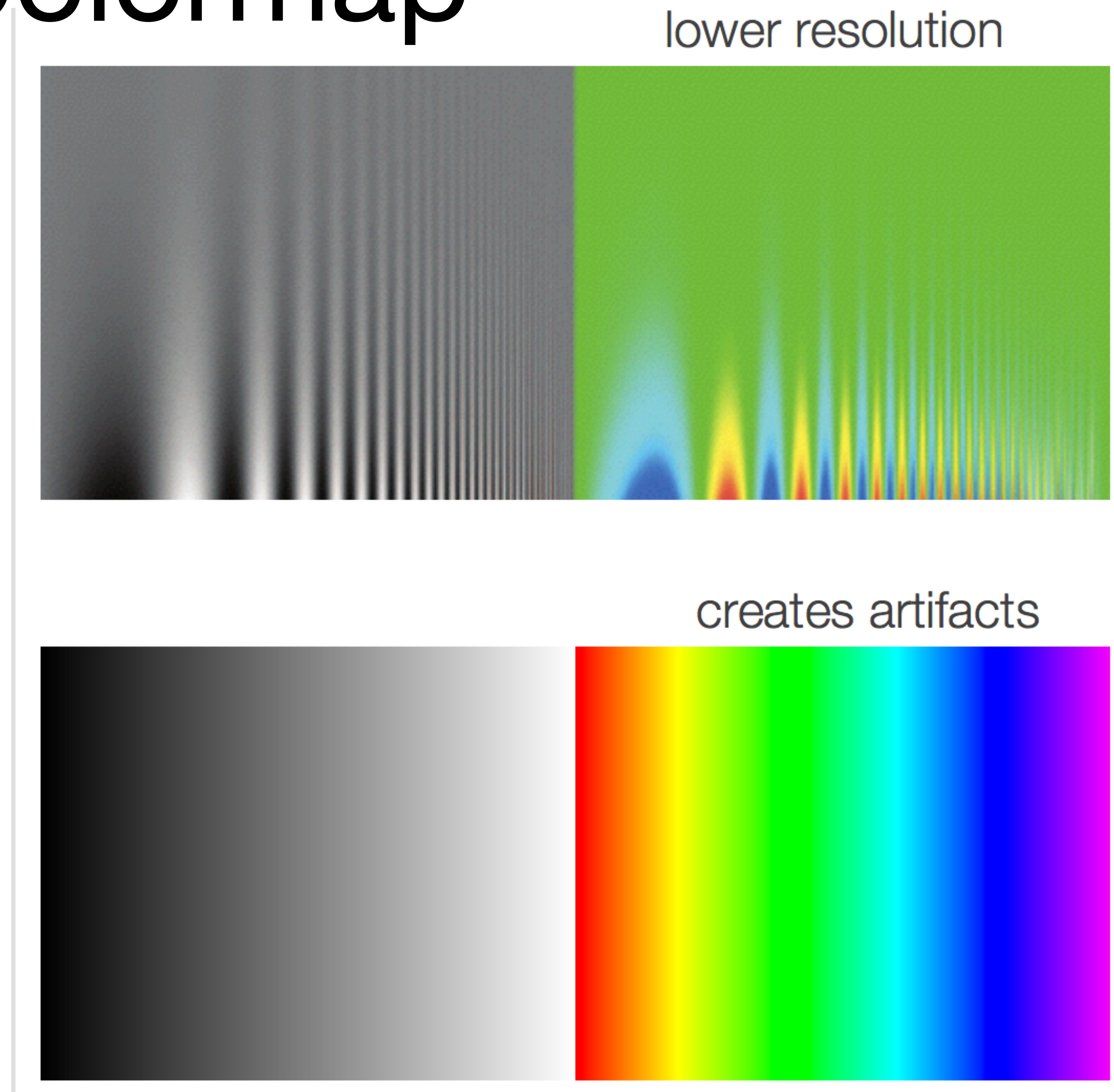
In this article, we reiterate the characteristics that mercials, weather forecasts, and even the IEEE Visualization Conference 2006 call for papers, just to name a few. The problem with this wide use of the rainbow color map is that research shows that it is rarely, if ever, the optimal color map for a given visualization.[1-6] Here we will discuss the rainbow color map's characteristics of confusing the viewer, obscuring data, and actively misleading interpretation.

Confusing

For all tasks that involve comparing relative values, the color map used should exhibit perceptual ordering. A simple example of a perceptually ordered color map is the gray-scale color map. Increasing luminance from black to white is a strong perceptual cue that indicates values mapped to darker shades of gray are lower in value than values mapped to lighter shades of gray. This mapping is natural and intuitive.

The Dreaded Rainbow Colormap

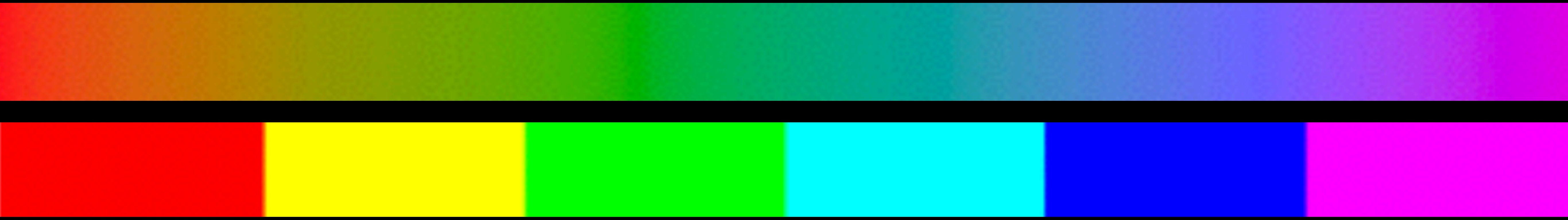




If you need going to use the rainbow colormap, use an **isoluminant** version, **quantize** it, or **both**

Bad

Better



Infovis 2011

Evaluation of Artery Visualizations for Heart Disease Diagnosis

Michelle A. Borkin, *Student Member, IEEE*, Krzysztof Z. Gajos, Amanda Peters, Dimitrios Mitsouras, Simone Melchionna, Frank J. Rybicki, Charles L. Feldman, & Hanspeter Pfister, *Senior Member, IEEE*

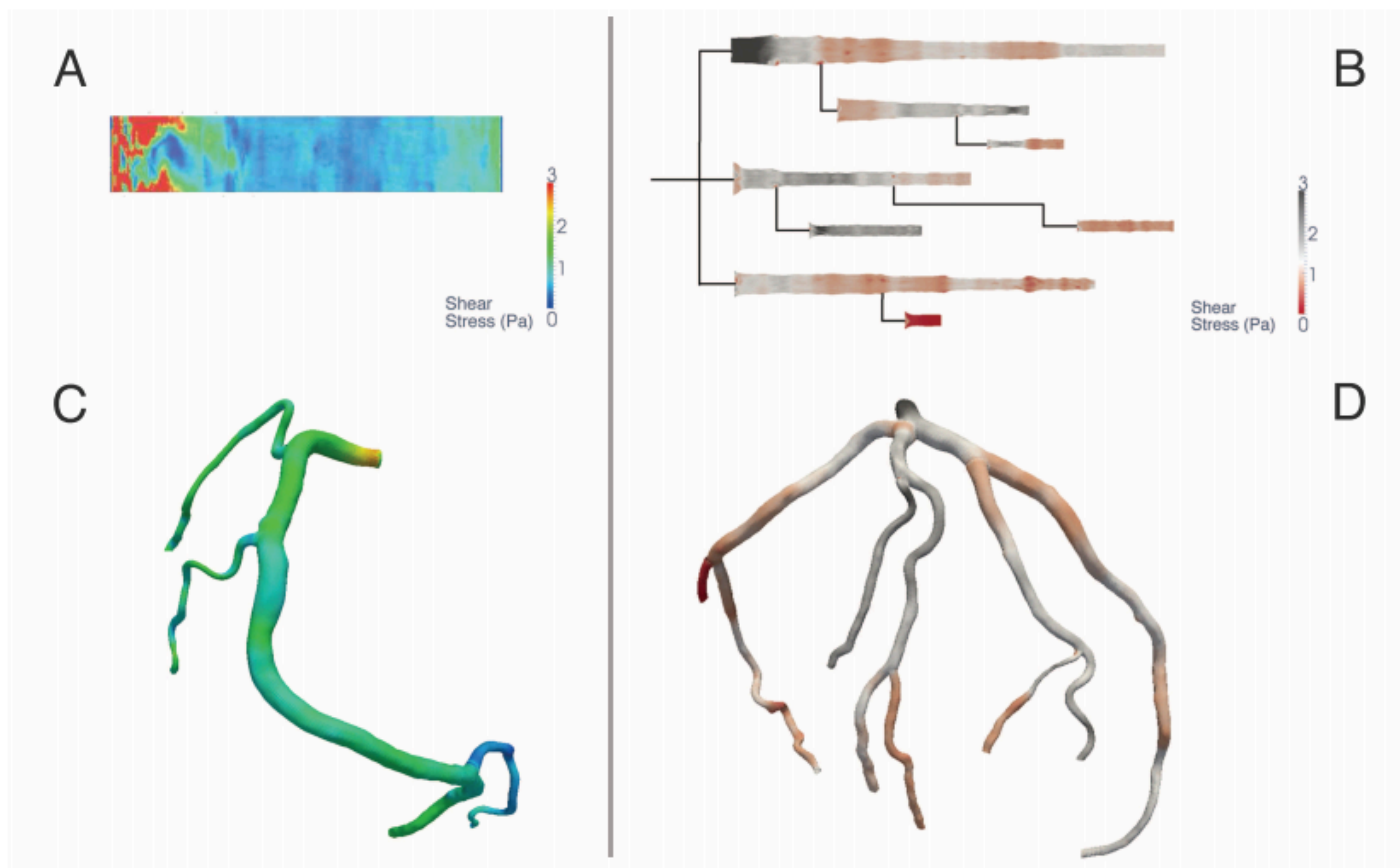
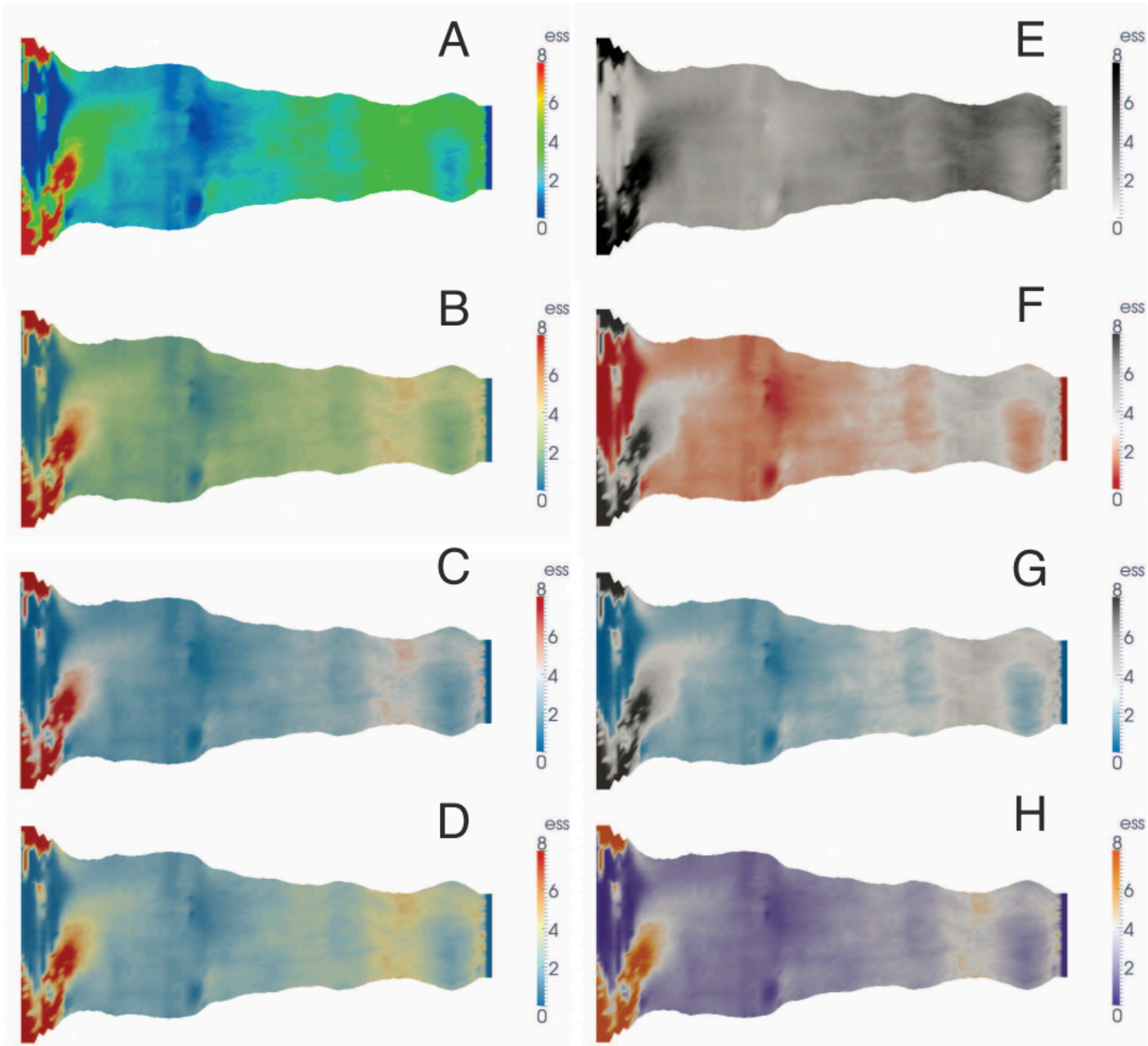


Fig. 1. Left: Traditional 2D projection (A) of a single artery, and 3D representation (C) of a right coronary artery tree with a rainbow color map. Right: 2D tree diagram representation (B) and equivalent 3D representation (D) of a left coronary artery tree with a diverging color map.

Abstract— Heart disease is the number one killer in the United States, and finding indicators of the disease at an early stage is critical for treatment and prevention. In this paper we evaluate visualization techniques that enable the diagnosis of coronary artery disease. A key physical quantity of medical interest is endothelial shear stress (ESS). Low ESS has been associated with sites of lesion formation and rapid progression of disease in the coronary arteries. Having effective visualizations of a patient's ESS data is

Borkin et al., Infovis 2011



Borkin et al., Infovis 2011

Colormap design matters very strongly

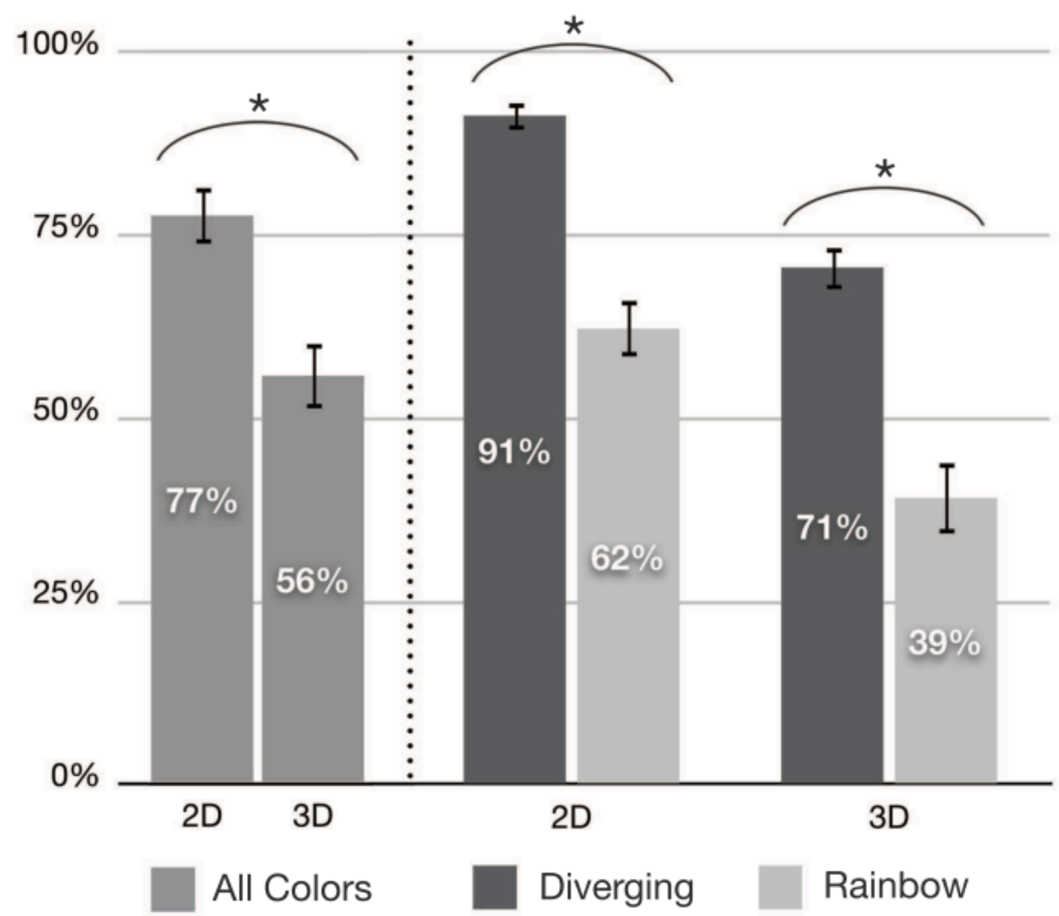


Fig. 7. **Average percent of low ESS regions identified** broken down by 2D and 3D representation, and color. Error bars correspond to the standard error and the asterisks indicate results of statistical significance. Participants were more accurate in 2D and when using the diverging color map.

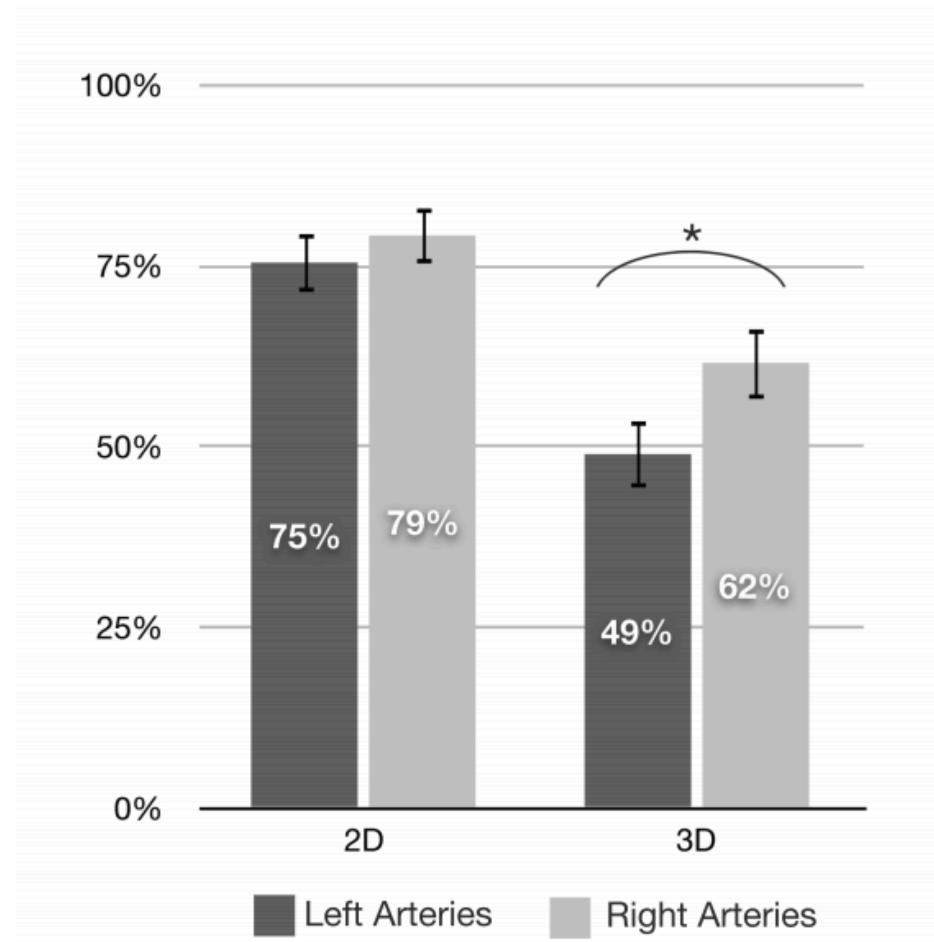


Fig. 8. **Average percent of low ESS regions identified** broken down by 2D and 3D representation, and left and right artery systems. Error bars correspond to the standard error and the asterisks indicate results of statistical significance. In 3D, users were less accurate identifying regions in the most complex data sets (i.e., left artery systems). Whereas in 2D, performance was the same regardless of task complexity.

COLORBREWER

COLORGORICAL