Assignment 5

Louis Rassaby, Susie Church, and Jim Bonish

24 November 2014

Contents

1 Introduction

An important part of data visualization is understanding how the presentation of data affects a viewer's perception of it. Bar charts, pie charts, and line charts all have merits, but tell different stories. Numerous studies have worked to evaluate different types of graphs on the accuracy of perception, but fewer have investigated the effects of colors on perception.

In this study, we aim to investigate the effects of color and movement on the perception of relative areas in bar charts.

2 Hypotheses

2.1 The Bright Color Hypothesis

Hypothesis: Marking two elements to be compared with solid bright green against a grey background will promote more accurate perception than marking those same elements against a white background.

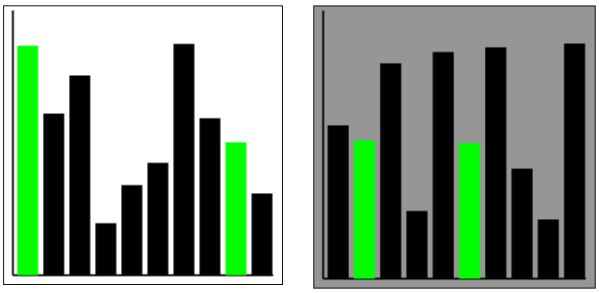


Figure 1: Bar charts on white or grey backgrounds

We hypothesize that the effect will be roughly proportional to the amount of grey on the screen in any particular graph. Therefore, the effectiveness of changing the background to grey will be most potent on the bar chart, somewhat less on the pie chart, and least effective on the treemap.

2.2 The Flashing Color Hypothesis

Hypothesis: Flashing between bright green and white at 1 second intervals will make perception less accurate than displaying elements in solid green.

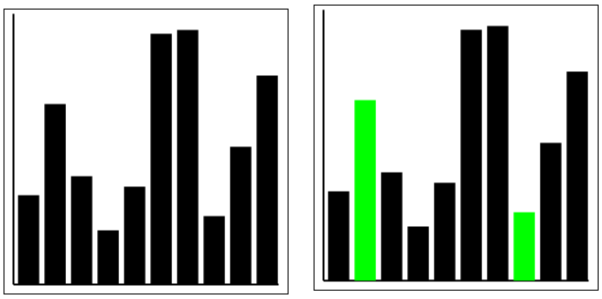


Figure 2: One bar chart in its *off* and *on* states

Flashing between white and green is distracting, but also draws the viewer's attention. We believe the distracting nature of flashing will impair the comparison abilities of viewers.

3 Methods

We evaluated 19 subjects, each on 20 tests, leading to a total of 380 observations. Though it was suggested that 30 could have been a better number of observations per subject, we chose to evaluate more subjects to account for fewer tests per subject.

Each of the 19 subjects viewed charts of only one type, broken up as follows:

1. 7 subjects on bars flashing between green and black at 1 second intervals against a white background

2. 6 subjects on solid green bars against a white background

3. 6 subjects on solid green bars against a grey background

On each test, participants were asked to enter what they thought was the ratio of the smaller of the two bars indicated compared to the larger. For example, if a viewer thought that the smaller bar were 3/4 the size of the larger, she may enter *75*, indicating that she thought the smaller bar was 75% the size of the larger.

4 Results

The error metric we used was the Cleveland and McGill's error. Out of all 380 observations, the mean for the error level was 1.427.

As indicated by *Figure 3*, the 95% confidence intervals were relatively wide. This is a factor of the small sample size of 19 subjects. In order to calculate confidence intervals, we assumed a normal distribution of errors and used the equation,

$$\text{Error Bound} = \mathbb{Z}_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

We plugged in $\alpha = 0.95$. This gave us a Z-score of $\mathbb{Z}_{\frac{\alpha}{2}} = 1.96$. σ is the standard deviation, and n is the number of trials. The results from this computation are in the "Error Bound" column of the following table, as well as in *Figure 3*.

Type	Mean	Standard Deviation	Error Bound at 95% CI
Flashing Bars on White Background	1.510	1.409	1.044
Solid Green Bars on Grey Background	1.528	1.664	1.331
Solid Green Bars on White Background	1.233	1.577	1.262

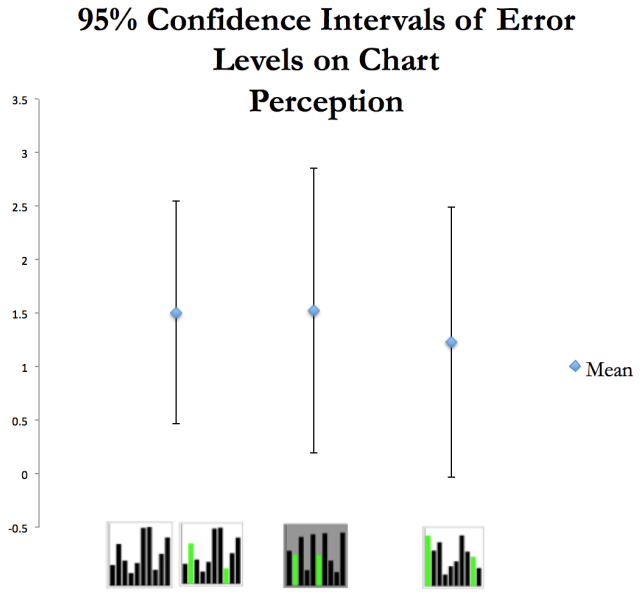


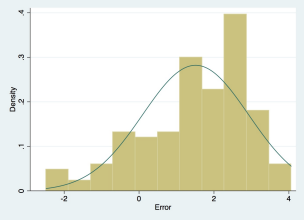
Figure 3: Bar charts (flashing, solid on grey, and solid on white) by accuracy of perception

4.1 Normality

4.1.1 Flashing Bars on White Background

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
Error	**140**	**0.94612**	**5.910**	**4.013**	**0.00003**

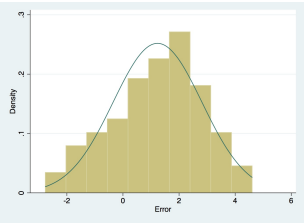


The Shapiro-Wilk test indicates that we can reject the null hypothesis that the error level is normally distributed at a 95% confidence interval. The data does look normally distributed, but the sample size is too small to know for sure.

4.1.2 Solid Green Bars on White Background

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
Error	**120**	**0.98069**	**1.858**	**1.388**	**0.08250**

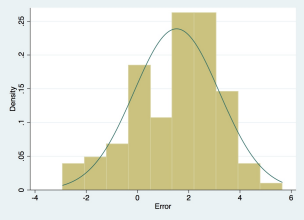


The Shapiro-Wilk test indicates that we cannot reject the null hypothesis that the error level is normally distributed at a 95% confidence interval. Therefore, it is likely that the data is normally distributed. It certainly looks normally distributed based on the graphs, but we can't get a better idea of the normality without more trials.

4.1.3 Solid Green Bars on Grey Background

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
Error	**120**	**0.96326**	**3.536**	**2.829**	**0.00233**



The Shapiro-Wilk test indicates that we can reject the null hypothesis that the error level is normally distributed at a 95% confidence interval. As with the first test, the data does look normally distributed, but the sample size is too small to know for certain.

4.2 Statistically Significant Differences in Population

As discussed in *Section 4.1*, it is impossible to know for certain, without further trials, whether the data is normally distributed. Given the information we have, we will therefore run the *Mann-Whitney U-test* on each of our hypotheses.

```
. ranksum Error if Vis != "FlashingBar", by(Vis)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      Vis |      obs       rank sum    expected
----------+---------------------------------
 GrayBar |      120          15405       14460
WhiteBar |      120          13515       14460
----------+---------------------------------
 combined |      240          28920       28920

unadjusted variance    289200.00
adjustment for ties         0.00
                       ----------
adjusted variance      289200.00

Ho: Error(Vis==GrayBar) = Error(Vis==WhiteBar)
           z =   1.757
    Prob > |z| =   0.0789
```

```
. ranksum Error if Vis != "GrayBar", by(Vis)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      Vis |      obs       rank sum    expected
----------+---------------------------------
FlashingBar |      140          19245       18270
   WhiteBar |      120          14685       15660
----------+---------------------------------
   combined |      260          33930       33930

unadjusted variance    365400.00
adjustment for ties         0.00
                       ----------
adjusted variance      365400.00

Ho: Error(Vis==FlashingBar) = Error(Vis==WhiteBar)
           z =   1.613
    Prob > |z| =   0.1068
```

Figure 4: *Mann-Whitney U-Test* results for grey v. white (left) and flashing v. white (right)

For the "Bright Color Hypothesis", our results indicated that the medians for each of the groups are not statistically different at any level smaller than 7.89%. We cannot reject the null hypothesis that there is no difference between graphs with green bars against a white or a grey background to a 95% confidence interval. Thus, our results are inconclusive.

For the Flashing Color Hypothesis, our results indicated that the medians for each of the groups are not statistically different at any level smaller than 10.68%. We cannot reject the null hypothesis that there is no difference between flashing bars and non-flashing bars to a 95%. Once again, our results are inconclusive.

5 Conclusions

Though the data seems to indicate that green bars against white background yield the best results, we can't say that this is true with a high degree of certainty without further trials.

It is quite likely that the error distributions are normal, particularly for the solid green bars against a white background, though this is also uncertain given the small sample size.

Though we have only scraped the surface of the perception of charts in this study, with more research, provable effects of color on barchart perception can be proven.