Emily Southern
INFO 4602/5602
Extra Credit Assignment

Part 1: Designing an Experiment

Write an experiment to test a specific question about an aspect of the effectiveness of visualizations. Note that this experiment should address a specific question that you can answer by measuring people's performance in a certain task.

In your writeup, include a concise but thorough description of the following aspects of the experiment:

For this experiment, I will investigate the effectiveness of different types of bar charts and color schemes for making comparisons. I will use data on the number of males and females from the top 10 represented countries at various marathons.

Research Question (What question will you address?)

How effective are different kinds of bar charts and color schemes for making comparisons?

Hypotheses (What do you think will happen and why?)

I hypothesize that paired bar charts will be more effective for comparing the number of males (M) and females (F) within a country and stacked bar charts will be more effective for comparing total number of runners from different countries. I also hypothesize that using an intuitive color scheme (pink for females/blue for males) will be more effective and using a counterintuitive color scheme (blue for females/pink for males) will be less effective.

Independent Variables (What will you change? What are the levels of each?)

I will manipulate the type of bar chart and the color scheme. There will be two levels of bar charts: paired bar chart and stacked bar chart. There will be two levels of color: pink for females/blue for males, and blue for females/pink for males.

Dependent Variables/Measures (What will you measure?)

Effectiveness of each type of visualization will be quantified by the number of correct responses to the following questions:

1) Were there more male or female runners from (insert country here)?
2) Were there more total runners from (Country A) or (Country B)?

Control Variables (What else do you need to account for? How will you do it?)

When comparing the number of correct responses for paired bar charts and stacked bar charts, I will need to account for the magnitude of difference between the number of males and females or the number of runners from each country. To do so, I will use the same dataset for each type of graph. I will also need to keep the size of each visualization and the width of the bars within each visualization constant. I will have the user use the same computer screen for the entire experiment so that the colors render the same for both types of bar charts. I will randomize the order of the visualizations to eliminate the effect of training or task familiarity on my dependent variable.

Description of the Stimuli (What will the participant see? Can be an annotated sketch or a verbal description. Include the source of any data needed)

The participant will be presented with either a paired bar chart or a stacked bar chart displaying the number of males and females from the top 10 represented countries of a specific major marathon. The color scheme of the visualization will either be intuitive (pink for females/blue for males) or counterintuitive (blue for females/pink for males). The data will be extracted from a website that reports the runners' demographics and finish times for more than 800 marathons each year (http://www.marathonguide.com).

Experimental Procedure (What will the participant do? Please describe this using a step-by-step procedure and include any details necessary to conduct the experiment)

The participant will open a program that initially prompts them for their name, age, and sex. The participant will click "Submit" and will be taken to a screen that displays information about the data that will be represented in the experiment. When ready to begin, the participant will click "Start" and will be taken to a screen displaying the first visualization. Each visualization will include a legend for the male/female color scheme in the upper right hand corner. The question will be displayed below the visualization, and the user will be able to select their answer from a choice of two radio buttons. Then, they will click "Next" and will be taken to a screen displaying the next visualization. Participants will be presented with a total of 80 visualizations: 10 datasets with 8 questions for each dataset, as described below:

For eat dataset…		Males vs. Females	Country A vs. Country B
Paired bar chart	Intuitive color scheme	Vis 1	Vis 5
	Counterintuitive color scheme	Vis 2	Vis 6
Stacked bar chart	Intuitive color scheme	Vis 3	Vis 7
	Counterintuitive color scheme	Vis 4	Vis 8

Planned Analysis (How will you analyze your dependent variables and why?)

To compare the effectiveness of paired bar charts and stacked bar charts, I will perform paired t-tests on the percentage of correct responses to questions for Vis 1 vs. Vis 3, Vis 2 vs. Vis 4, Vis 5 vs. Vis 7, and Vis 6 vs. Vis 8. To compare the effectiveness of intuitive and counterintuitive color schemes, I will perform paired t-tests on the percentage of correct responses to questions for Vis 1 vs. Vis 2, Vis 3 vs. Vis 4, Vis 5 vs. Vis 6, and Vis 7 vs. Vis 8. By making these specific comparisons, I am able to isolate the effect of each independent variable to be sure that I know what is causing the (potential) difference in percentage of correct responses.