

**1. What is our independent variable? What is our dependent variable?**

The independent variable is the word shown, specifically whether it matches the color of the word. The dependent variable is how fast participants can identify the color of the word.

**2. What is an appropriate hypothesis for this task? What kind of statistical test do you expect to perform? Justify your choices.**

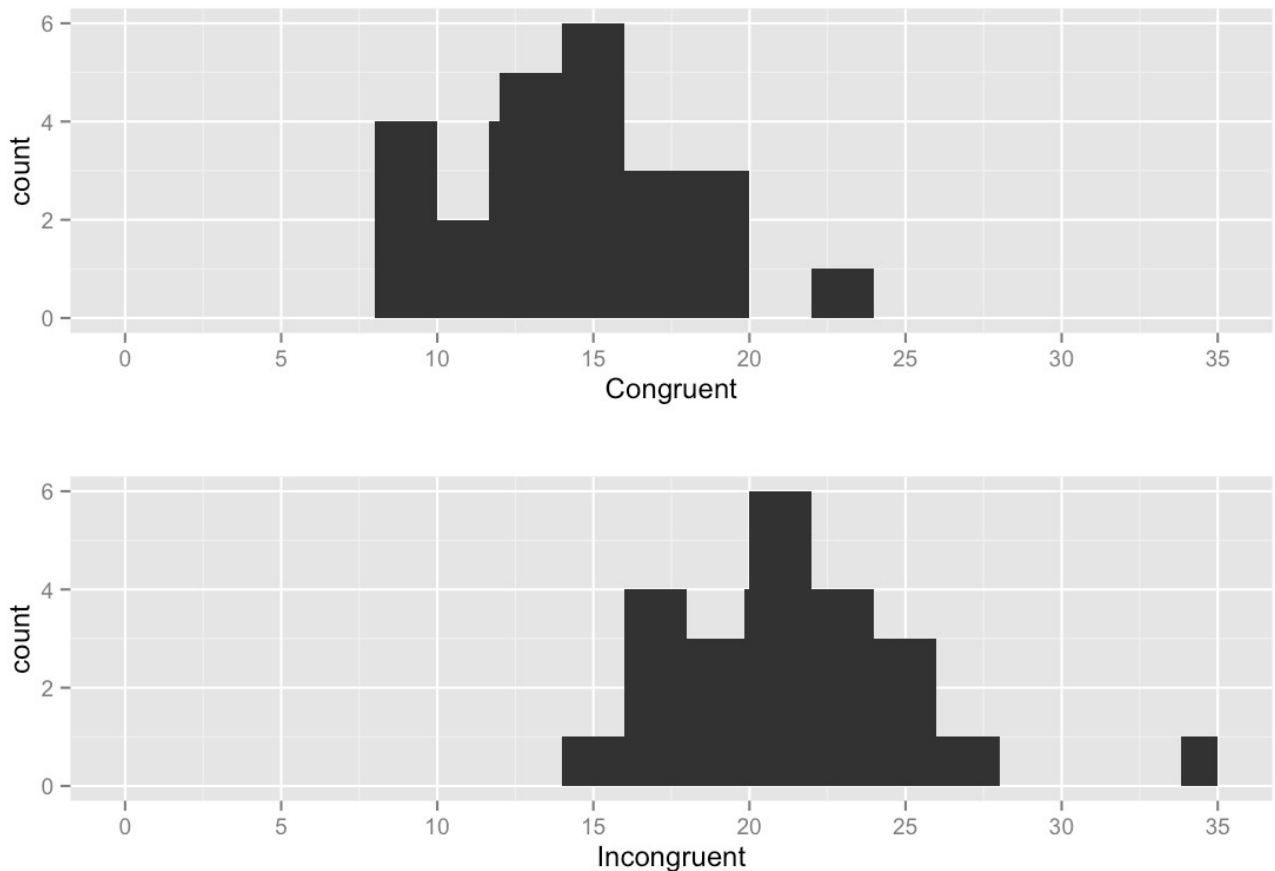
The null hypothesis would be that the congruency does not affect performance or affects it negatively ( $\mu_1 \leq \mu_2$ ). I am guessing that congruent words will be faster to identify than incongruent words, so the alternative hypothesis would be that the congruency affects performance positively. I am hypothesizing this because I believe that people will be more used to reading words on paper than identifying colors, and so it will be more difficult with incongruent words. To analyze the results, I will perform a dependent sample one-tailed t-test.

**3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.**

The average time for congruent words was ~14.051, with a sample standard deviation of ~3.559. The average time for incongruent words was ~22.016, with a sample standard deviation of ~4.797.

**(go to next page)**

**4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.**



The above plot shows histograms of the distributions of congruent and incongruent times. The Incongruent plot appears to be further to the right than the congruent plot, showing that the distribution is higher.

**5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?**

Running the t-test yields a t-statistic of ~8.021. This t-statistic is incredibly high. Because this statistic is so high, I can use a t-critical value of 3.768 for a confidence level of 0.9995. Therefore, I can easily reject the null hypothesis and say that if a word is congruent the time to identify the color will most likely be shorter. This agrees with my original prediction. However, I was surprised how high the t-statistic was. With an r-squared value of nearly 0.74, the congruency is the single biggest factor affecting the outcomes, eclipsing any individual variation for other reasons.

**6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!**

I believe that the reason for these results is tied to reading. We are so used to reading any words that we

see that we automatically read them if we see them. It is hard for us to disregard the words that we see to focus purely on color. A similar example could be done with size or font instead of color (eg. 'SMALL', 'medium', 'big').