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

The independent variable is the two lists of colored words (congruent and incongruent). They are constant throughout the experiment for all participants.

The dependent variable is the measured time to speak-out the colors.

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

In the Stroop effect experiment each participant has to perform two tasks with different conditions: one task with a congruent and another task with an incongruent word list. This results in paired samples with two conditions and therefore a dependent t test is appropriate.

The null hypotheses for the test is that the time to speak out the colors does not change wether a congruent or an incongruent list of words is used.

$H_0 : \mu_C = \mu_I$ or $\mu_D = \mu_C - \mu_I = 0$

The alternative hypothesis is that it makes a difference if the word lists are congruent or incongruent. As I expect that the times for an congruent list are shorter, I use a one tailed test

$H_A: \mu_C < \mu_I$ or $\mu_D = \mu_C - \mu_I < 0$

I use a one tailed dependent t test for paired samples to test if there is a significant increase in the times needed to speak out the colors or not.

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

Sample size: 24
Degrees of freedom: 23

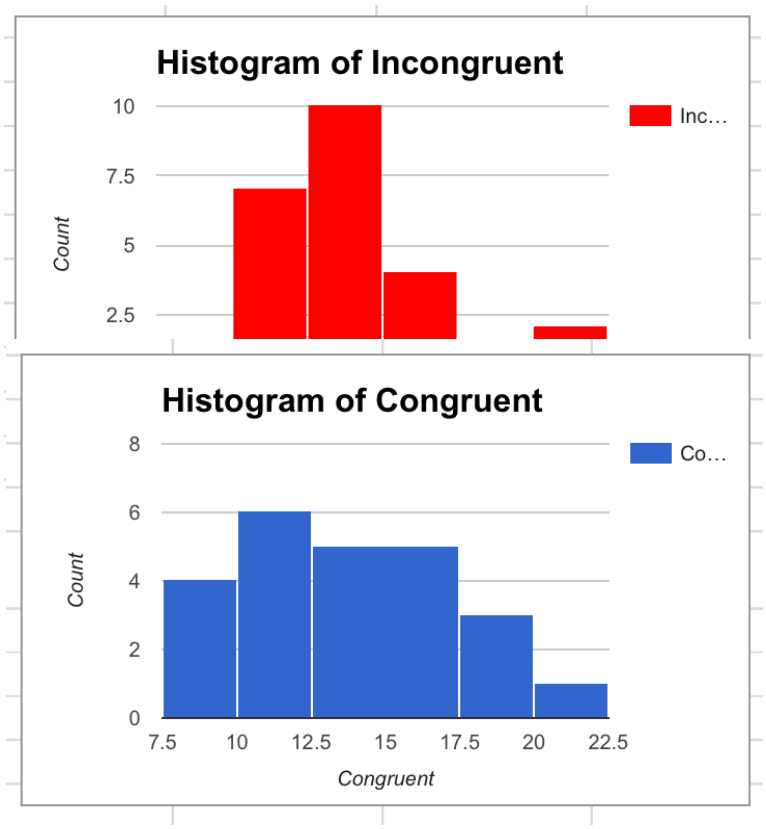
Mean for congruent word list: $\mu_C = 14.05$
Mean for incongruent word list: $\mu_I = 22.02$

Standard deviation for congruent word list: $SD_C = 3.56$

Standard deviation for incongruent word list: $SD_I = 4.80$

Mean for the difference: $\mu_D = -7.96$
Standard deviation for the difference: $SD_D = 4.86$

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 sample for the congruent word list is distributed over 7.5 to 22.5 seconds with the mode between 10 and 12.5 seconds.

The sample of the incongruent word list is distributed over 12 to 36 seconds with the mode between 20 and 24 seconds.

Where the distribution of the congruent sample is relatively flat the distribution of the incongruent sample shows a peak and some values far out.

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?

The t-statistics for this test is $t = \mu_D / (SD_D / \text{sqrt}(25)) = -8.02$

The p-value for this t-statistics is less than 0.0001 so the confidence level is more than 99.99%.

Even for α-level of 0.001 the critical t-statistics is 3.768 (or -3.768).

Therefore I reject H_0

The time needed to speak out the congruent word list is significant less than the time needed to speak out the incongruent word list.

This matches my expectation because it is irritating if the color of a word is not what the word means.

NB: as a non-native speaker the effect is not that large for me. It is larger if the words are written in my language (German).

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 think the effect is due to the irritating difference between the color of the word and the meaning of the word itself. We read the words as „pictures“ and „identify“ the color before even seeing the color itself.

This should also work with pictures and words and there are children's songs that play with a similar effect when combining rhymes with words.

A. used material

t-table : https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg
p value calculation: http://www.graphpad.com/quickcalcs