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

The **independent variable** are the printed words (color names) in a color of ink. The color names are either matching the font color (congruent) or not (incongruent). The **dependent variable** is the seconds to name the colors of ink of the printed words.

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

The **null hypothesis (H_0)** could be that the mean time that it takes to name the ink color for incongruent words or congruent words is not different. With other words the mean difference is zero. μ_i represents the mean for incongruent words and μ_c the mean for congruent words.

$H_0: \mu_i = \mu_c \rightarrow H_0: \mu_i - \mu_c = 0$

The **alternative hypothesis (H_A)** could be that mean time to name the ink color for incongruent words to be bigger than for congruent words. I would expect that because of the conflicting signals if the color name and the actual color of the word are different, it would take longer to name the ink color for incongruent words.

$H_A: \mu_i > \mu_c$

I would expect to perform a **dependent t-test for paired samples**. The same group of subjects is used for reading the lists with the congruent and the incongruent words. Each subject is in both samples. So we compare the difference mean for the same subjects. We assume that the differences are approximately normally distributed and this is a random sample of the population.[12]

[1] Own notes from Udacity "Statistics" classes.
[2] McClave, J.T., Benson, P.G., and Sincich,T. *Statistics for Business and Economics.* Upper Saddle River, Prentice-Hall, Inc., 2001, p.412-419.

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

There are 24 participants in the study ($n_D = 24$). For this sample of participants the **sample mean difference** ($\bar{x}_D = \bar{x}_i - \bar{x}_c$) is 7.96. Meaning that on average the participants needed 7.96 more seconds to read the list with incongruent words. Most participants needed between 5 to 10 seconds more to read the incongruent words (**mode**, see histogram below) and the **median** was 7.67 seconds. As the median is smaller than the mean we can conclude the distribution is positively skewed.

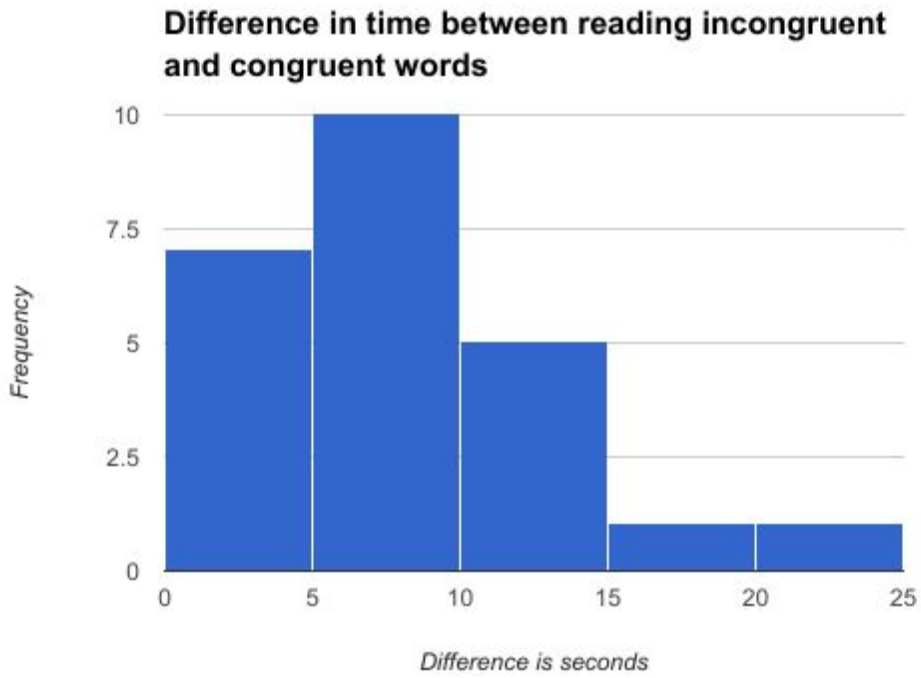
The more time needed to read the incongruence words **ranges** between 1.95 and 21.92 seconds. The **sample standard deviation of differences** (S_D) is 4.86 seconds. The standard deviation measures the variance of a dataset from its mean. As we assume that our data set is normally distributed, approximately 68% of the data should fall within one S_D of the mean and approximately 95% within two S_D of the mean.[3,4]

Calculations can be found in this Google spreadsheet.

[3] Own notes from Udacity "Statistics" classes.
[4] McClave, J.T., Benson, P.G., and Sincich,T. *Statistics for Business and Economics.* Upper Saddle River, Prentice-Hall, Inc., 2001, p.54-59, 65-69.

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 data is positively skewed as can be seen in the histogram below. Most participants (22) needed less than 15 seconds more to read the incongruent words. Two participants needed more time and one can be regarded as an outlier (Q3 + 1.5 * IQR = 20.88).[5]



Calculations can be found in this Google spreadsheet.

[5] See also here for calculations of IQR and outliers.

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?

$t(23) = 1.714, p < .05$, one-tailed test

Test static is 8.02 ($p = .00$[6]). As the test statistic falls in the rejection region we reject the null hypothesis, we can conclude (at $\alpha = .05$) that the difference in the mean seconds needed to read the words differs from 0. The correlation coefficient (r^2) is .74, meaning that 74% of the difference in means can be explained by the difference in the printed words (matching or not matching the font color).

On average participants need 7.96 more seconds to read the incongruent words than the congruent words which is what I expected (95% CI: (5,91 , 10.02)). My own results for doing the Stroop tasks was 13.80 seconds.

Calculations can be found in this Google spreadsheet.

[6] The p-value is .0001, but as the APA style suggest to round at 2 decimals I have used .00.

Sources

- Own notes of Udacity “Statistics” classes.
- Wikipedia page on Stroop effect, https://en.wikipedia.org/wiki/Stroop_effect.
- McClave, J.T., Benson, P.G., and Sincich,T. *Statistics for Business and Economics*. Upper Saddle River, Prentice-Hall, Inc., 2001.
    - p. 54-59: Measures of central tendency.
    - p. 65-69: Measures of variability.
    - p. 412-419: Paired difference experiments.

Calculations can be found in this Google spreadsheet.