

The Stroop Experiment

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

In the Stroop experiment, the independent variable is the words condition, either congruent or incongruent.

The dependent variable is the time taken by the participant at the experiment to name the ink 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.

The null hypothesis for this experiment is that the change of words condition does not affect the time it takes to name the ink colors, or said differently that the average time taken to name the ink colors is similar in both congruent and incongruent words conditions. The alternative hypothesis is that the average time taken to name the ink colors is different in the two words conditions.

In symbolic signs, that gives:

$$H_0: \mu_C = \mu_I$$

$$H_a: \mu_C \neq \mu_I$$

with μ_C the average time taken to name the ink colors in the congruent words conditions and μ_I the average time taken to name the ink colors in the incongruent words conditions.

Knowing that the tow samples are dependent and the size of our sample (24 participants) is smaller than 30, we will perform a two-tailed t-test with to reject or retain the null hypothesis. We could also have performed a one-tailed t-test, because it seems intuitive that the incongruent words condition is increasing the time taken by the participants to name the ink colors (we expect $\mu_C < \mu_I$). However, we have chosen to do a two-tailed test in case our intuition is not correct and $\mu_C > \mu_I$.

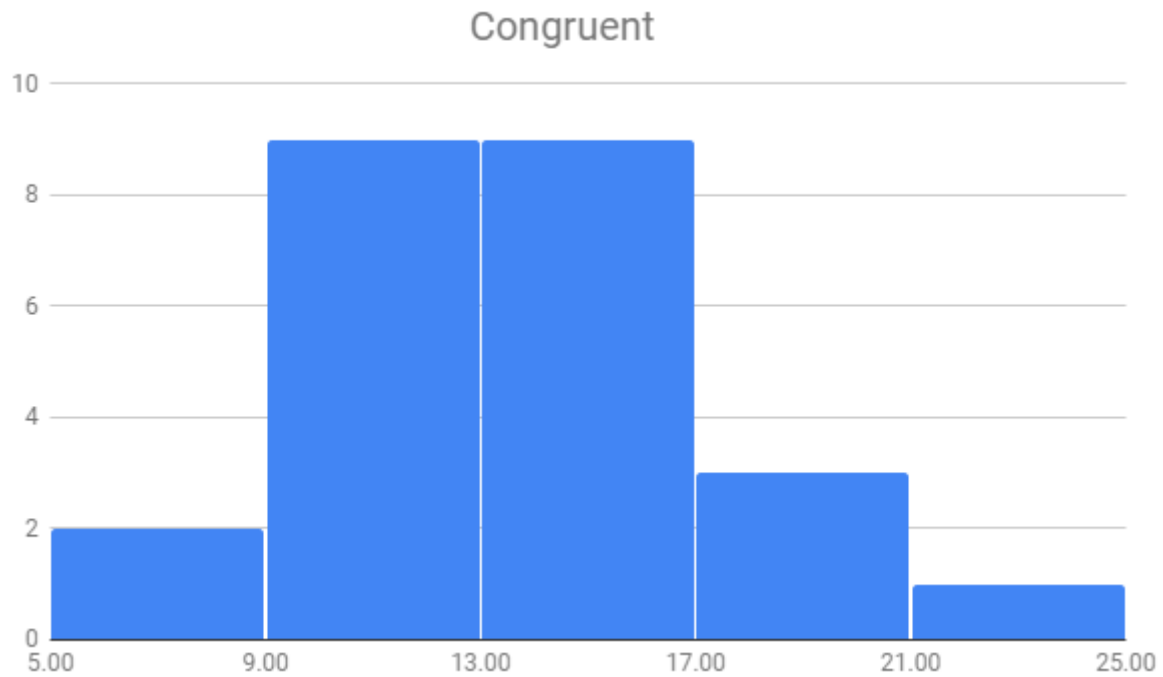
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

The following measures describe our sample dataset:

	Congruent	Incongruent	Difference (Congruent – Incongruent)
Average	14.051	22.016	- 7.965
Variance	12.669	23.012	23.667
Standard deviation	3.559	4.797	4.865

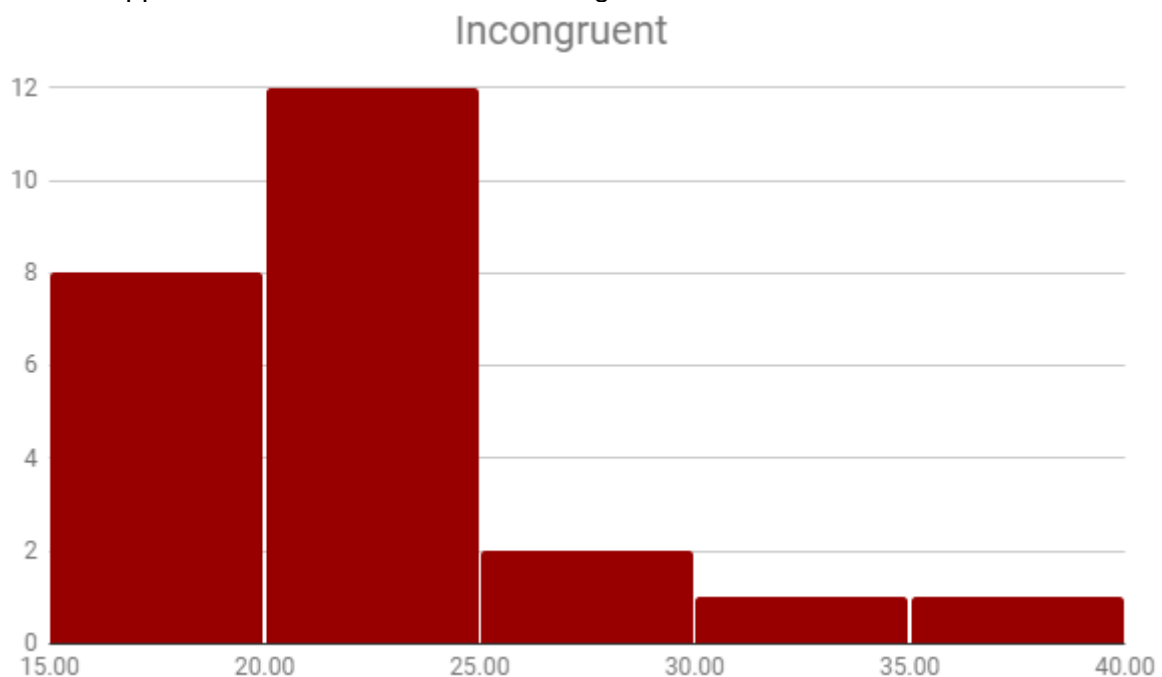
4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.

Here below is displayed an histogram showing the distribution of the congruent data (bucket size: 5):



It is quite clear that the distribution is approaching a normal distribution shape, but not yet there due to the small size of the sample. We can also notice that as expected the distribution is more or less centered around the average, 14.

The same applies to the distribution of the incongruent data:



The distribution is also approximately centered around the average, 22, and approaching a normal distribution shape.

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?

We choose a high confidence level: 99.9%, meaning $\alpha = 0.001$.
With 23 degrees of freedom, that leads to $t_{critical} = \pm 3.768$.

The statistical results are summarized in below table:

Standard Error of the Mean SEM	0.993
$t_{statistics}$	- 8.021
p-value	<0.0001
Cohen's d	- 1.637
r^2	0.737

We have $t_{statistics} < t_{critical}$, we can therefore reject the null hypothesis because the difference between the two sample means is statistically significant at 99.9%. Or said differently, the time taken by a participant at a Stroop experiment to name the ink color is greater on average for the incongruent words condition.

r^2 being equal to 0.737, we also know that 73.7% of the difference between the two means are statistically due to the words condition (or dependent variable).

These results match up with our intuition, that the incongruent words condition make it way harder to recognize the ink color. That is probably because in this case the brain has to process two contradictory information, the word and the ink color.