

# Investigation of a Stroop Test dataset

## Background Information

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the color of the ink in which the word is printed. The task has two conditions: a congruent words condition, and an incongruent words condition. In the congruent words condition, the words being displayed are color words whose names match the colors in which they are printed: for example RED, BLUE. In the incongruent words condition, the words displayed are color words whose names do not match the colors in which they are printed: for example PURPLE, ORANGE. In each case, we measure the time it takes to name the ink colors in equally-sized lists. Each participant will go through and record a time from each condition.

## Questions for Investigation

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

-> Our Independent variable is the stroop test conditions and the Dependent variable is the time taken by the subjects to complete naming the ink colors in seconds.

### 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 appropriate Hypotheses set for the given task would be :

$$H_o : \mu_I \leq \mu_C \quad \text{OR} \quad \mu_C - \mu_I \geq 0$$

$$H_a : \mu_I > \mu_C \quad \text{OR} \quad \mu_C - \mu_I < 0$$

Here  $\mu_C$  is the average time taken to name the ink colors for the Congruent conditions and  $\mu_I$  is the average time taken to name the colors for the Incongruent condition in seconds

- The above null Hypothesis states that the time taken to name the ink colors for the Incongruent condition is not significantly longer than that of the time taken to name the ink colors for the Congruent condition.
- The above alternative hypothesis states that the time taken to name the ink colors for the Incongruent condition is significantly longer than that of the time taken to name the ink colors for the Congruent condition.

-> As the population parameters for the given sample are unknown and both the treatments are carried out on the same subjects hence we should perform a one sample paired 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.**

-> Measures of central tendency :

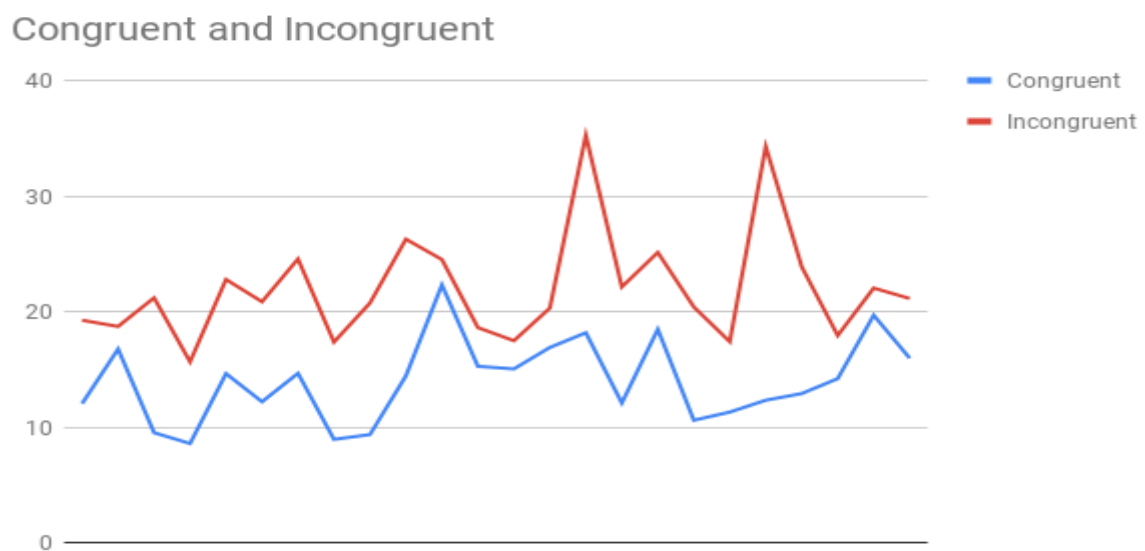
- Average time taken to name congruent ink colors :  $\mu_C = 14.051$
- Average time taken to name Incongruent ink colors :  $\mu_I = 22.016$
- Average of the differences for time taken for the two conditions :  $\mu_{C-I} = -7.965$

-> Measures of Variability :

- Standard deviation of the time taken in the Congruent condition :  $S_C = 3.559$
- Standard deviation of the time taken in the Incongruent condition :  $S_I = 4.797$
- Standard deviation of the difference of the time taken in both conditions :  $S = 4.865$

**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.**

-> The line chart visualization for the sample data would be :



(The Y axis represents the time taken to name the ink colors)

From the above visualization, It's clear that the time taken to name the ink colors in the Incongruent condition is always greater than the time taken to name the ink colors in the Congruent condition.

**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 have :

- Sample size (n) = 24
- Mean of the differences ( $\mu_C - \mu_I$  OR  $\mu_{C-I}$ ) = -7.965
- Standard deviation of the differences of the time taken (S) = 4.865

- Confidence level ( $\alpha$ ) = 95%
- Degrees of freedom (df) = 23
- Critical statistic value ( $t^*$ ) = 1.714

$$\begin{aligned}\text{Hence } t \text{ statistic } (t) &= (\mu_c - \mu_i)/(S/\sqrt{n}) \\ &= -7.852\end{aligned}$$

-> **Conclusion :**

Here t statistic (t) value is far away from the t critical ( $t^*$ ) value, Hence :

$$p < 0.05$$

and we reject the null hypothesis , Which means that the time taken to name the ink colors in Incongruent condition is signifacntly longer than the time taken to name the ink color in Congruent condition .

-> The results do match up with my expectations as it was clear from the line chart of the data that for almost each subject the time taken to name the ink colors in the Incongruent condition was longer than the time taken in the Congruent condition.

**6. 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!**

-> There are a few theories like : Selective attention, Autoaticity, parallel distributed processing etc. Which offer possible explanation to the effects of the stroop test<sup>[1]</sup> (link in the references).

-> Dimesional overlap<sup>[2]</sup>, Comparing color-word and picture-word<sup>[3]</sup> would also result in effects similar to the stroop test( link in the references).

-> Researchers also use the Stroop effect during brain imaging studies to investigate regions of the brain that are involved in planning, decision-making, and managing real-world interference (e.g., texting and driving).

## References :

1. [https://en.wikipedia.org/wiki/Stroop\\_effect](https://en.wikipedia.org/wiki/Stroop_effect)
2. <http://dimensional-overlap.com/the-stroop-like-task/>
3. <https://www.ncbi.nlm.nih.gov/pubmed/8090863>