

Data Analyst Nanodegree Project: Testing a Perceptual Phenomenon

By: Geeta Dandamudi

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.

Question 1:

What is our independent variable? What is our dependent variable?

Dependent Variables: The time taken (by the participant) in responding to what the color of the printed word is or the response time.

Independent Variables: The word in the list, irrespective of it being the name of the color itself or a different word or the word being congruent or incongruent.

Question 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 Hypothesis for the stroop effect:

Null Hypothesis: The mean time taken to recognize the words from the congruent list should be equal to or greater than the mean time taken to recognize the words in the incongruent

Alternative Hypothesis: The mean time taken to recognize the words from the congruent list is less than the mean time taken to recognize the words from the incongruent list.

The Hypothesis can be represented as:

$H_0: \mu_C \geq \mu_I$

$H_A: \mu_C < \mu_I$

Where, H_0 is the Null hypothesis

H_A is the Alternate hypothesis

μ is the Population mean

μ_C is the Mean time taken to read a word from the congruent list

μ_I is the Mean time taken to read a word from the incongruent list.

Statistical Test to perform:

Based on the hypothesis a one-tailed dependent samples t-test must be performed comparing the differences in the means for the congruent and incongruent word lists.

Reason for the chosen statistical test:

As the alternate hypothesis is that the mean time for the incongruent words is less than the mean time for the congruent words, a one- tailed test in the negative directions is apt for it.

we need to base the results for the entire population on the sample data alone as the entire population statistics aren't available.

The two groups compared here are dependent and are based on same subject, so a t-test must be performed to analyze it.

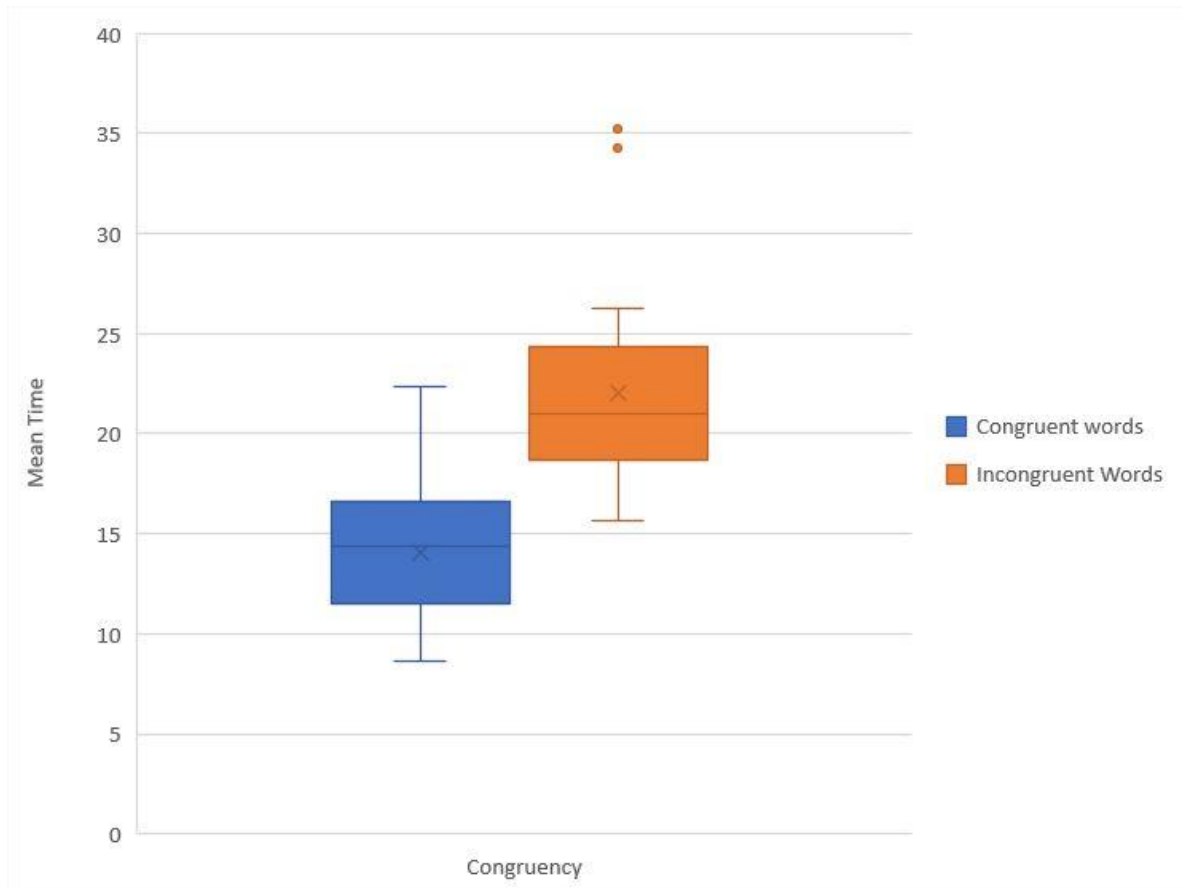
Question 3:

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

Statistic	Congruent	Incongruent	Difference
Mean	14.05	22.02	-7.97
Median	14.36	21.02	-6.66
Variance	12.67	23.01	23.67
Standard Deviation	3.56	4.8	4.86
Standard Error	0.73	0.98	0.99

Question 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 box plot clearly shows the significant difference in mean times of the two word groups: Congruent and incongruent. The two words have considerably different ranges.

Question 5:

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?

$$\alpha = .05$$

$$df = 23$$

$$t_{\text{critical}} = -1.714$$

$$t = -8.02$$

$$p\text{-value is } < .0001$$

There is sufficient evidence at $\alpha = .05$, the t critical value is -1.714 and the calculated T value with the difference in mean times of congruent and incongruent lists is -8.02, which is far below the critical value.

Since P value is less than the negative t-value, it is a one tailed t-test in negative direction i.e. the mean time for recognizing the congruent words is less than the time taken to recognize the incongruent words. Hence, the evidence is sufficient to reject the null hypothesis (The times taken to recognize the words in both the lists is not equal)

Question 6:

What do you think is responsible for the effects observed?

In my opinion, this effect is caused due to the distraction by the words being displayed. The participant gets distracted by the word and this interferes in the response time for recognizing the color of the word.

The interference between the different information our brain receives **causes** a problem, the interference occurs because words are read faster than colors are named.

Several variations to the stroop effect can be found online like the : Warped words, emotional, spatial, numerical and reverse stroop effects.

References:

<https://faculty.washington.edu/chudler/words.html>

https://en.wikipedia.org/wiki/Stroop_effect

<http://www.thesimpledollar.com/the-stroop-effect-and-your-wallet/>

