Test a Perceptual Phenomena

By Eric Fraser

Question 1: Identify the variables in the experiment

In this experiment, our dependent variable is the amount of time it takes to complete the test, and the independent variable is the layout of the colors and the words associated with the colors, whether they are congruent or incongruent.

Question 2a: Establish Hypotheses

Note that i = incongruent data, c = congruent data, and μ represents the population mean.

Null Hypothesis:

$$H_0: (\mu_i - \mu_c) = 0$$

The average difference in time it will take the population to complete the incongruent color/word test compared to the congruent color/word test is statistically 0.

Alternative Hypothesis:

$$H_a: (\mu_i - \mu_c) > 0$$

The average difference in time it will take the population to complete the incongruent color/word test compared to the congruent color/word test will be statistically significant, with the incongruent test taking more time than the congruent test.

Question 2b: Establish a statistical test

There are several reasons for using a t-test as opposed to a z-test or f-test in this experiment. Because we are testing for the difference in mean between two variables from the same population, the sample size is small, and we do not know the population standard deviation, we will be using a paired T-Test for our analysis to compare congruent and incongruent test scores in an effort to measure the Stroop Effect. A z-test, like the t-test is useful for applying statistics to normal distributions. A z-test would be used if the population standard deviation was known, or if we were dealing with a larger sample size. Additionally, the sample observations are dependent on each other, so a t-test would make more sense than a z-test here. An f -test would be used if we were trying to compare the variance between two

populations, which is not relevant to this particular experiment. Assumptions made by this test are that both sample sets are from the same population, and also that all other variables are controlled for except for the time it takes to complete the test.

Question 3: Report descriptive statistics

Measures of centrality chosen for this experiment are the median and mean of the samples for both congruent subjects and incongruent subjects. The most useful measure of variability would be standard deviation, the results of which calculated for each set of the samples can be found below. Note that because we are using a sample size and not a population, our standard deviation measurement considers degrees of freedom to be n-1 and not n.

Central Tendency	Congruent	Incongruent
Mean	14.05	22.02
Median	14.36	21.02

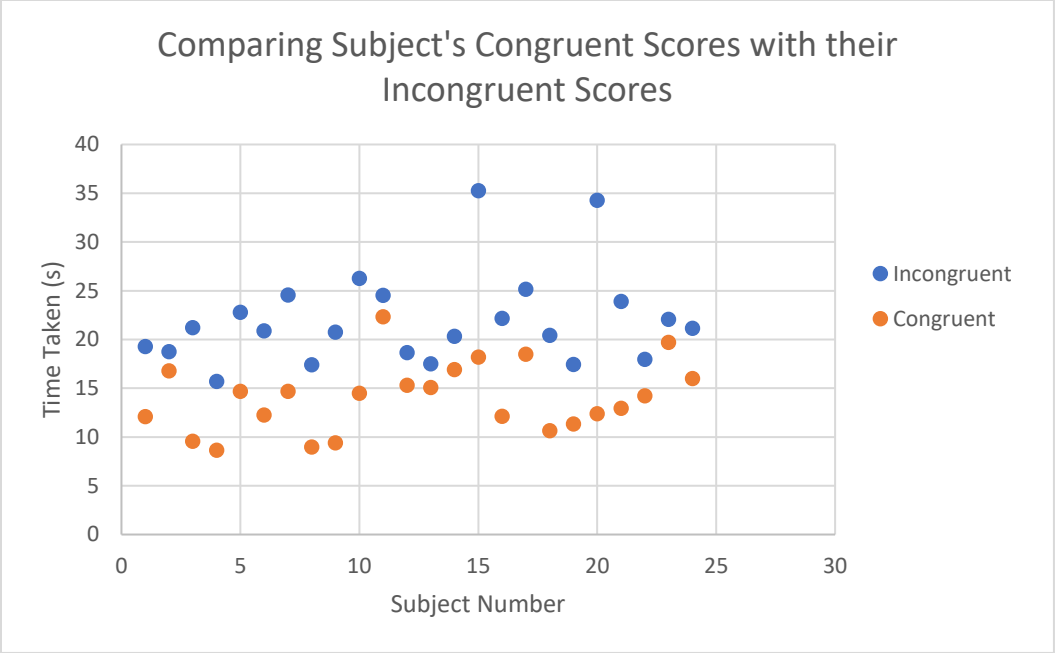
Variability	Congruent	Incongruent
Standard Deviation	3.56	4.80

Question 4: Plot the data

Here is the table of our test results. Below you can see a scatterplot graphing both sample's time taken for each test by subject number.

Congruent	Incongruent
12.079	19.278
16.791	18.741
9.564	21.214
8.63	15.687
14.669	22.803
12.238	20.878
14.692	24.572
8.987	17.394
9.401	20.762
14.48	26.282
22.328	24.524
15.298	18.644
15.073	17.51
16.929	20.33
18.2	35.255
12.13	22.158
18.495	25.139
10.639	20.429
11.344	17.425
12.369	34.288
12.944	23.894
14.233	17.96
19.71	22.058
16.004	21.157

You can see from looking at the graph that every single person in the sample took more time to complete the incongruent test than it took them to complete the congruent test. You can also see that variability in incongruent scores appears higher – this is confirmed above using the standard deviation for both sets of data. Interestingly, while each person's time taken increased, there appears to be significant variability in individual aptitude for the test, meaning there are several examples of subject's incongruent times beating the congruent times of their peers.



Question 5: Perform the statistical test and interpret your results

To perform the T-Test, we must first gather some descriptive statistics regarding the differences between incongruent and congruent scores.

If you subtract each incongruent score by its congruent pairing, then take the mean of that set, the average difference comes out to:

Δ_{mean} = 7.96 seconds

Taking the standard deviation of the differences, we get the value:

Δ_s = 4.86

So the question becomes: If H_0 is true and there is, on average, no difference between congruent and incongruent scores, how likely is it that we got an average sample difference of 7.96 seconds?

To find this out, we need to use the formula below:

$$test = \frac{\Delta \text{mean}}{\Delta \text{s} \,/\, \sqrt{n-1}}$$

$$test = \frac{7.96}{4.86 \,/\, \sqrt{24-1}}$$

Calculating the above, we get a test value of 7.85. This tells us that our sample average is about 7.85 standard deviations above what we would expect if we assume H_0 is true. With a t value of 7.85 and 23 degrees of freedom, the p-value is going to be less than 0.0001.

Based on the above, If H_0 is true, there is a statistically miniscule chance of seeing an average difference in values of 7.96 seconds. Therefore, we can reject our null hypothesis. *We have evidence to believe that our alternative hypothesis, that using incongruent color-word pairs in the Stroop Effect experiment will increase the time it takes to complete the test, is correct.*

Sources

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

https://faculty.washington.edu/chudler/java/timestc.html

https://www.le.ac.uk/oerresources/ssds/numeracyskills/page_19.htm