

P1: TESTING A PERCEPTUAL PHENOMENON

(STROOP TEST)

1. INDEPENDENT & DEPENDENT VARIABLES:

In our case, the independent variables are names and the colors of the words displayed, and the name/color congruency.

Whereas, the dependent variables are: one - the time (measured in seconds) taken by individuals to complete the congruent section, and two - the time taken to complete the incongruent section.

2. CHOICE OF HYPOTHESIS & STATISTICAL TESTS:

The **null hypothesis (H_0)** for this test should be that there is no significant difference in the population means of the time taken by people to read out all the given words when the words being displayed are in the same color as its name (congruent), and when the words being displayed are of a color different than that of its name (incongruent).

$$M_I - M_C = 0^1$$

Going by the differences in the sample means and medians, and my personal experience taking the test, I have a hunch that it should take participants longer to read out the words when the words are incongruent than when they are congruent.

Therefore, the **alternative hypothesis (H_1)** should be that the population mean of the time taken by individuals to read out the words when their names and colors match (congruent) is significantly greater than the time taken when the names and colors do match (congruent).

$$M_I - M_C > 0$$

Since we do not have the knowledge of population parameters, we will have to use a t-test to test our hypothesis. This choice can be justified because the assumptions for the t-tests are met.

One, we have a small sample size (less than 30 samples)

Two, we do not have any information on the population standard deviation.

¹ M_I is the population mean of the incongruent section, and M_C is the population mean of the congruent section.

And, given the nature of treatment and the alternative hypothesis, it is prudent that we perform a **one tail (dependent) paired t-test** because the same sample is afforded two different treatments: all individuals in the sample have to first answer the congruent and then the incongruent section; and we want to test whether there is increase in the time taken to complete the second section. This means that the independent variable is categorical, but it's categories are paired, thus meeting with the first assumption of a dependent t-test.

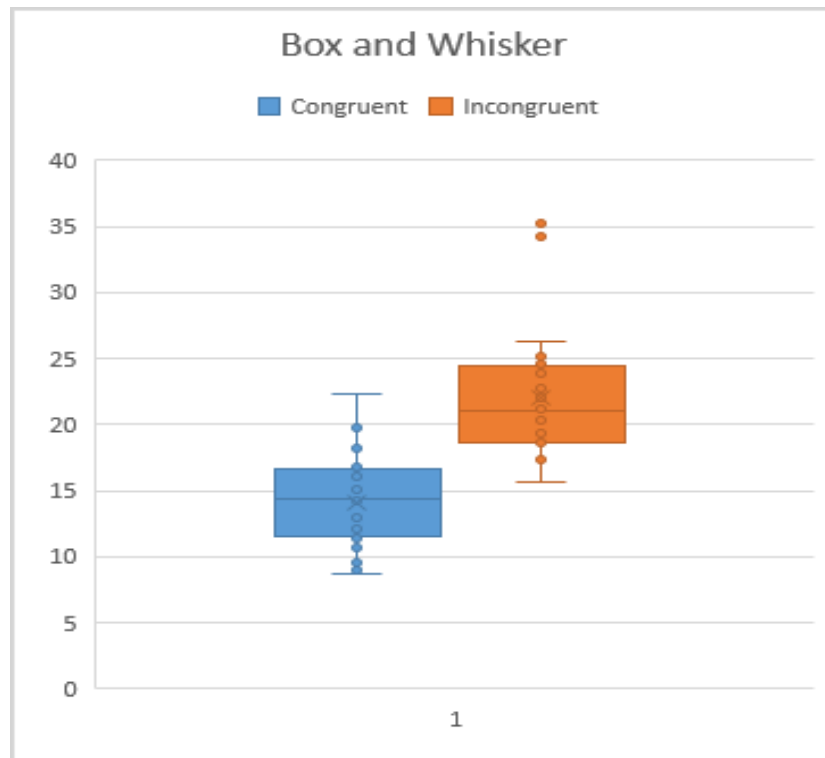
Our dataset is also consistent with the second underlying assumption of a dependent sample t-test. The differences between the dependent variables are not normally distributed. Even though these differences are significant at a level of 0.05, it's not so exaggerated for us not to use the t-test.

3. DESCRIPTIVE STATISTICS:

| MEASURES OF CENTRAL TENDENCY | | |
|-------------------------------------|------------------|--------------------|
| | Congruent | Incongruent |
| Mean | 14.051125 | 22.0159167 |
| Median | 14.3565 | 21.0175 |
| Mode | 12.079 | 19.278 |
| DISPERSION | | |
| Minimum | 8.63 | 15.687 |
| Maximum | 22.328 | 35.255 |
| Range | 13.698 | 19.568 |
| Q1 | 11.9 | 18.72 |
| Q3 | 16.2 | 24.05 |
| IQR | 4.3 | 5.33 |
| Variance | 12.669029 | 23.011757 |
| Std Deviation² | 3.559358 | 4.79705712 |

4. VISUALISATIONS:

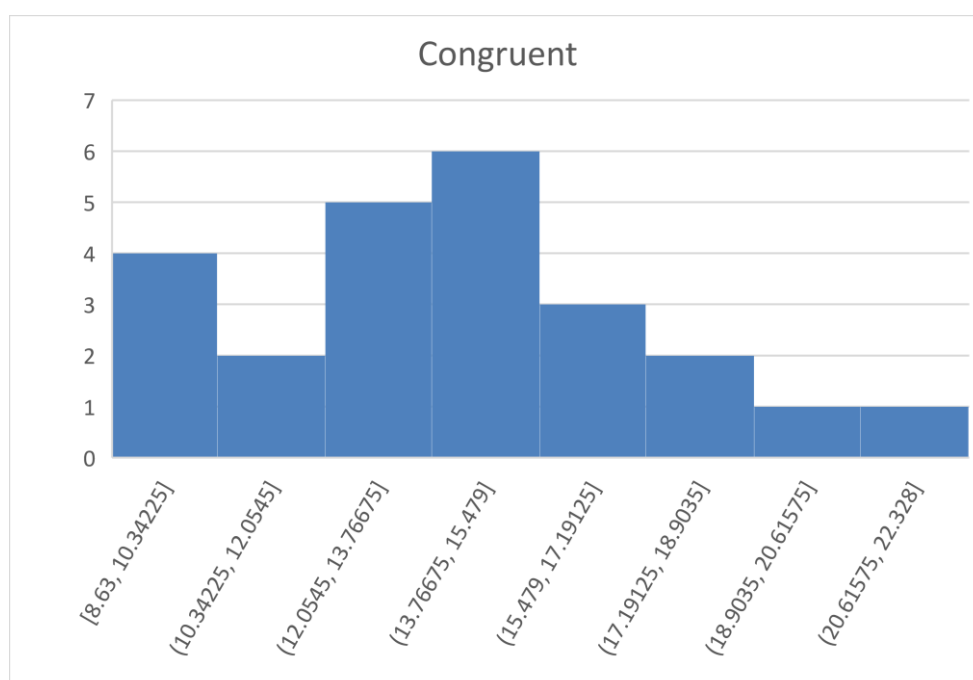
² This is the sample standard deviation and includes Bessel's correction.



The box and whisker plot gives us an idea of how the data is spread. We can see from the visualisation that there are **two outliers** on the incongruent section – two people who took unusually long to answer – with times of 35.25 seconds and 34.288 seconds.

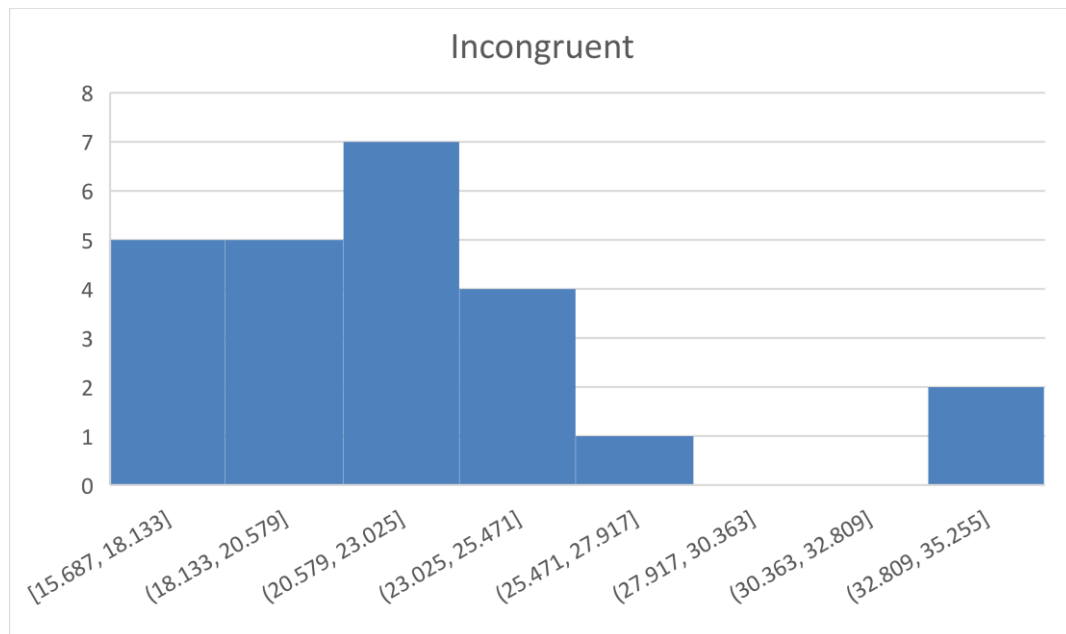
CONGRUENT SECTION:

The scores in this section has a range of **13.698** seconds, with a minimum value of **8.63** seconds and a maximum value of **22.328** seconds. The histogram reveals that the data has a slight **positive skew**.



INCONGRUENT SECTION:

The scores in this section range from a minimum value of **15.687** seconds to a maximum value of **35.255** seconds. This gives that section a range of **19.568** seconds. The histogram below, is **positively skewed**.



5. RESULTS

Paired Sample Statistics

| | Mean | N | Standard Deviation | Standard Error of Mean |
|-------------|------------|----|--------------------|------------------------|
| Congruent | 14.051125 | 24 | 3.559358 | 0.726550901 |
| Incongruent | 22.0159167 | 24 | 4.797057 | 0.979195185 |

Paired Sample Test

| | Paired Differences | | | | | | | |
|-------------------------|--------------------|--------------------|----------------|---|------------|----------|----|--------------------|
| | Mean | Standard Deviation | Standard Error | 95% Confidence Interval of the Difference | | t | df | Significance Level |
| | | | | Lower | Upper | | | |
| Incongruent - Congruent | 7.964792 | 4.864827 | 0.993029 | 6.26274059 | 9.66684275 | 8.020707 | 23 | .05 |

The confidence level is 95%

The paired sample t-test indicated that the time taken to complete the incongruent section ($M = 22.02$, $SD = 4.80$) was significantly higher than the time taken to complete the congruent section ($M = 14.05$, $SD = 3.56$), $t(23) = 8.02$, $p < .0001$, one-tailed.

Since $p < 0.05$, it means the p-value falls in the critical region, and is statistically significant. Therefore, the null hypothesis H_0 should be rejected.

We can conclude that people take anywhere between 6.26 seconds to 9.67 seconds longer on average to complete the incongruent section than the congruent section.

6. OPTIONAL:

I think the difference in times is due to conflicting stimuli in the second part: the text reads a color different from its font color. This means the brain will have to dedicate more time and attention to process the stimulus.

On reading up on the Stroop Effect, I discovered that originally the test was designed to have three stimuli – neutral, congruent and incongruent. However, the test administered as a part of this project takes into account only two of these stimuli. Congruent, where the ink color and the text refer to the same color. And, incongruent – where the ink color and the text refer to different colors.

The differences in times are observed due to interference of stimuli.

I found three theories explaining this effect, one of them being ‘Selective Attention’.

Alternative tasks of the sort include a numerical stroop test that underscores a relationship between numbers and physical sizes.