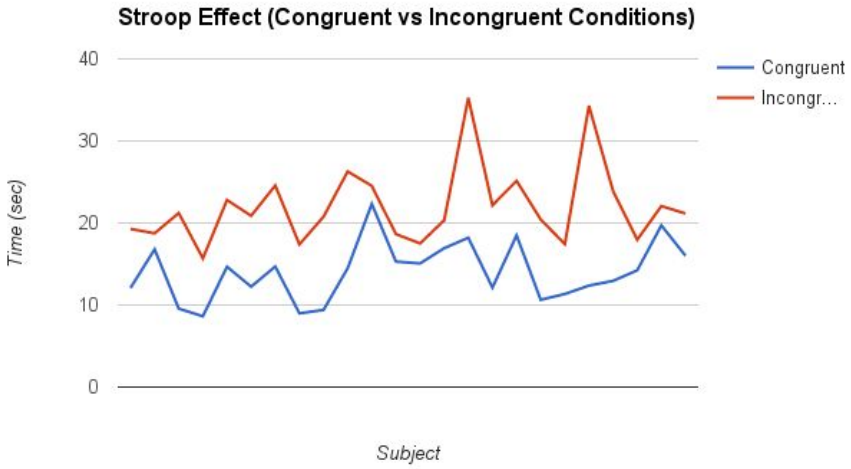
Udacity DAND Project 1: Stroop Effect
Nathaniel Booth

Variables
Independent variable: word condition, either congruent or incongruent.
Dependent variable: performance time (in seconds) elapsed to complete each condition.

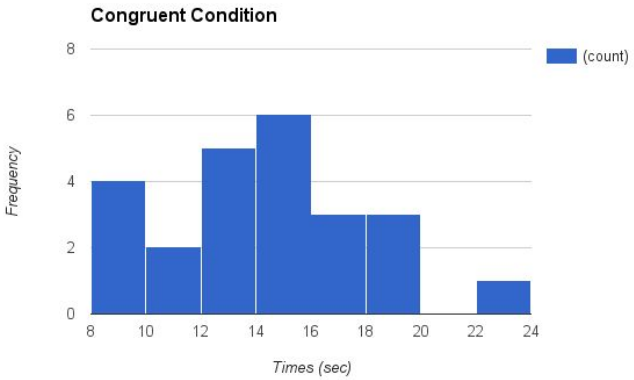
Hypotheses
Null hypothesis: the performance time will not be affected by the word condition.
Alternative hypothesis: the performance time will be significant decreased under the incongruent word condition.

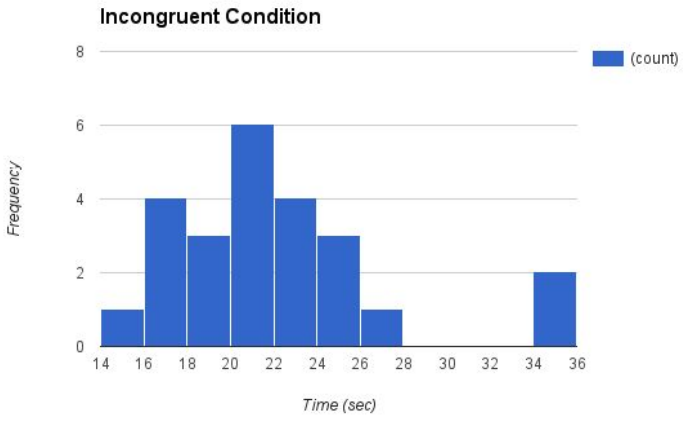
Test
It is given that each subject will perform under both conditions and the population parameters are unknown. Therefore, these are paired dependent samples in a within-subject design that will require a t-test.

Sample Data
From the chart at the right (Stroop Effect), the subjects in incongruent test condition appear to have significantly higher performance times compared to the same subjects under the congruent test condition.



The following two histograms (Congruent Condition and Incongruent Condition) suggest roughly normal distributions in performance time within each test condition with a couple outliers in the incongruent case.





The means for each condition are about 14 seconds and 22 seconds, respectively. The standard deviations are about 3.6 and 4.8, respectively. The higher variability in the incongruent condition is expected due to the presence of two apparent outliers.

Results

T-statistic	-8.02
P-value[1]	<0.0001
Alpha level	0.01
T-critical value[2]	-2.5
99% Confidence interval	(-10.45, -5.48)

The absolute value of the t-statistic is greater than the absolute value of the t-critical value, the p-value suggests that this outcome was likely not due to chance, and the t-statistic falls within the expected range with 99% confidence. The null hypothesis is rejected. The performance time of the subjects under the congruent condition is significantly lower than their performance under the incongruent condition.

Conclusion

This result is expected from a perusal of the sample data and consistent with the intuition that human brains will slow down in order to resolve the conflicting information input of the observed color and the printed color name.

Four theories are common[3], but parallel distributed processing theory most resonates with me. The conscious recognition of words is used and strengthened more often than the conscious recognition of colors. Therefore, the brain's initial impulse is to consider the word before the color. The conscious then recalls the goal of the task, inhibits the initial response, and falls back to the weaker processing pathway. The opportunity cost of processing the word before the color explains the increased processing time in the task.

A similar effect is expected any time the brain must rely upon a weak processing pathway to overcome a dominant processing pathway. The numerical Stroop[4] is an example, which experiments with the relationship between numbers and physical sizes associated with them. For example, the number 4 is expected to have a larger physical size than the number 3. If 3 is physically larger than 4, then the brain requires additional time to use the weaker conscious processing pathway for responding to physical size.

[1] The p-value is calculated from the t-statistic and 23 degrees of freedom using http://www.graphpad.com/quickcalcs/pValue1/.
[2] The t-critical value was found using the t-table at https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg.
[3] https://en.wikipedia.org/wiki/Stroop_effect#Theories
[4] https://en.wikipedia.org/wiki/Numerical_Stroop_effect