

Andrew Burruss
alburruss@gmail.com

Project 1: Stroop Experiment

Responses to Questions For Investigation

1. Variables

The independent variable is the colored word chart which each test subject read. We refer to the two charts as the Congruent, and Incongruent data sets. The dependent variable is the response time of the subject (for both Congruent and Incongruent data), measured in seconds.

2. Hypotheses and statistical tests

There were 24 test participants (to be referred to as the “sample”) taken from a larger population. Each participant gave a response time (measured in seconds) for both Congruent and Incongruent Stroop tests. Because of the small sample size and the design such that each data pair was generated by the same participant, we choose to use the paired t-test statistic. A priori, we know the results of other Stroop experiments (<https://faculty.washington.edu/chudler/words.html#seffect>), and we expect that this sample will behave in the same manner. That is, we expect the average response times for the Incongruent test to be greater than that of the Congruent test. This assumption about the sample justifies the use of a right one-tailed t-test, and we will assume a 0.01 significance level. That is to say, because of our expectation for this sample, we can focus our attention on the likelihood of the average Incongruent response time being greater than the average Congruent response time, and ignore the possibility that, vice versa, the average Congruent response time may be greater than the average Incongruent response time.

To conduct the paired t-test for this sample, we need to construct a variable which measures the difference between Incongruent and Congruent test response times in each data pair. Let X be the difference of the Incongruent response time minus the Congruent response time for each subject in the sample, measured in seconds.

$$X = (\text{Incongruent response time}) - (\text{Congruent response time}).$$

We use the symbol \bar{X} to represent the mean value of X for this sample. Statistical analysis will indicate if the sample statistic, \bar{X} , may be used to make a valid inference about Stroop experiment response times in the population from where this sample of participants was taken. Let U be the average difference of the Incongruent response time minus the Congruent response time for the population. We then formulate our hypotheses with respect to U . The null hypothesis represents the case where our t-test statistic indicates that the data for this sample does not provide sufficient statistical evidence that the average Incongruent response time should be significantly greater than the average Congruent response time in the population. Therefore, let the null hypothesis be that the value of U in the population is equal to or less than zero.

Null Hypothesis: $H_0: U \leq 0$

The alternate hypothesis represents the claim that this sample provides sufficient statistical evidence that the average Incongruent response time is greater than the average Congruent response time in the population. Let the alternate hypothesis be that the value of U in the population is greater than zero.

Alternate Hypothesis: $H_1: U > 0$

We will find the t-statistic for \bar{X} in this sample, its corresponding p-value, and a 99% confidence interval estimate of U . (Please see results and analysis in section 5.)

The t-statistic will quantize how likely it may be that the differences between the Incongruent and Congruent response times for this sample can be explained as a random occurrence. The decision making process for the hypotheses is conducted as follows. If the t-test indicates that the value of \bar{X} is less than 1% likely to have occurred randomly for this sample, then we will reject the null hypothesis and accept the alternate hypothesis as a statistically significant inference about the population. If the t-test indicates that the value of \bar{X} for this sample is a number which can be explained as a random occurrence for 99% of possible samples of this size, then we will fail to reject the null hypothesis. That is, the sample would not provide statistically significant evidence that U should be greater than zero.

The p-value measures the probability of the t-statistic to occur randomly for this sample. The confidence interval uses the t-test statistic and the value of \bar{X} in this sample to estimate an interval where we expect, with 99% likelihood, to find the true value of U in the population.

3. Descriptive statistics

Paired samples: X = Incongruent response – Congruent response

$N = 24$

Mean = 7.96479

Variance = 23.66654

Standard Deviation = 4.86448

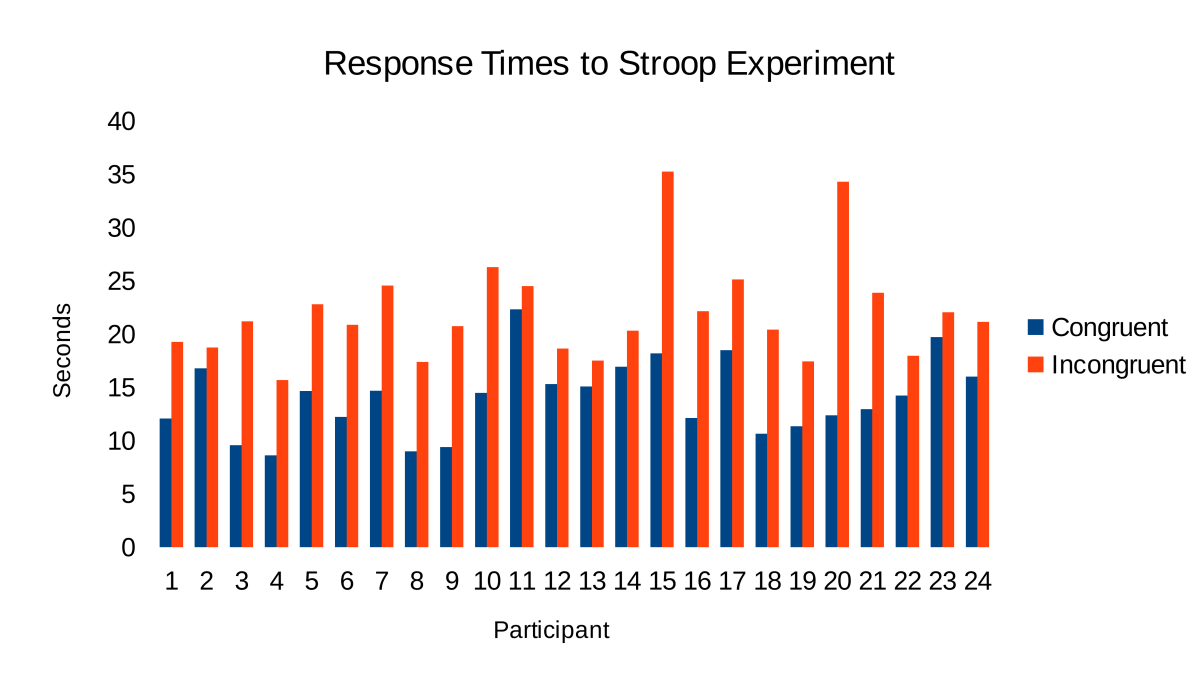
Min = 1.95

Max = 21.919

Range = 19.969

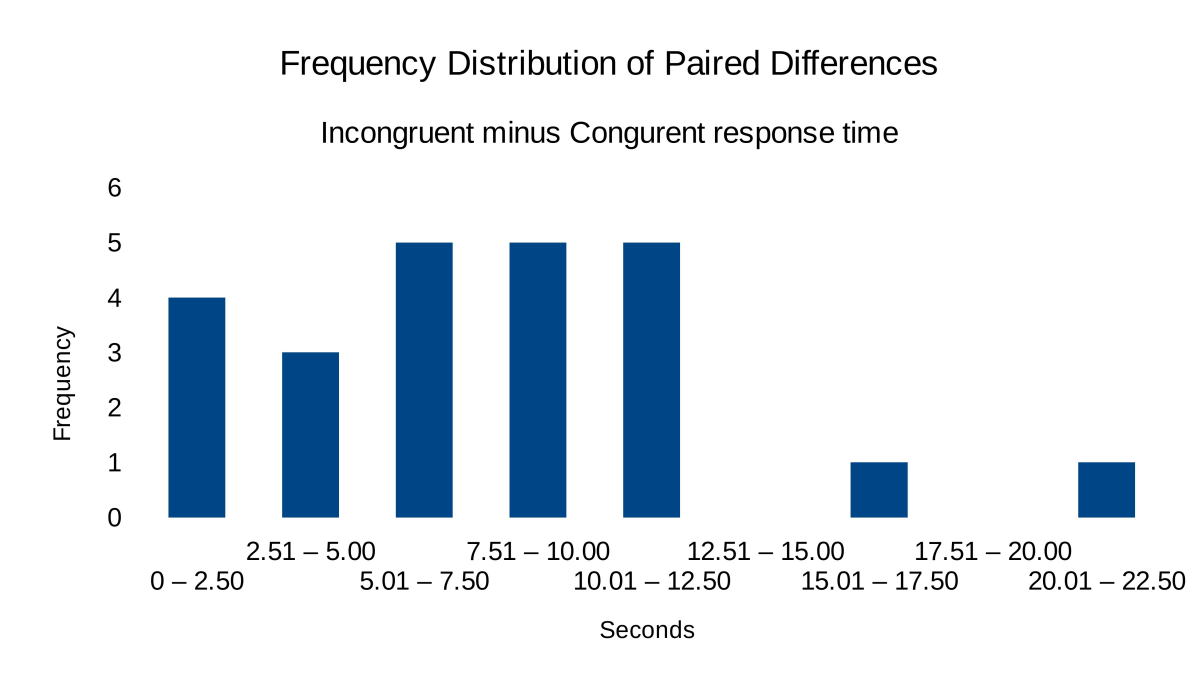
4. Visualizations

The paired response times of each Stroop experiment participant is diagrammed below.



We see that every response is greater where the test participant reads the Incongruent values, compared to the Congruent values. This agrees with the a priori expectation of Incongruent response times to be greater. In particular, participants #15 and #20 have the greatest increases of response time when reading the Incongruent values, and participants #12 and #13 have the least increases of response time. Note that this diagram shows *both* the responses for each participant. We will perform statistical analysis on the differences of the paired responses (X , as defined above).

The frequency distribution of the differences of the paired data samples (X) is diagrammed below.



Each value diagrammed is the Incongruent response minus the Congruent response, measured in seconds, for each pair. The frequencies are grouped in widths of 2.50 seconds. There were no differences less than 0 or greater than 22.00 seconds. There are three modal groups; five responses occur in each of the 5.01 – 7.50, 7.51 – 10.00, and 10.01 – 12.50 (sec) groups. The mean of 7.96479 appears in the second modal group. Moreover, the 0 – 2.50 (sec) group has a frequency of 4, and the 2.51 – 5.00 (sec) group has a frequency of 3. There are two individual outliers in the 15.01 – 17.50 and 20.01 – 22.50 groups, respectively. This indicates that the data set is approximately uniformly distributed in the range of 0 to 12.50 seconds.

5. Statistical test and report

Each value X (measured in seconds) is the difference within each data pair of

$$X = (\text{Incongruent response time}) - (\text{Congruent response time}).$$

\bar{X} = mean value of X in the sample

U = average value of X in the population

Null Hypothesis

$$H_0: U \leq 0$$

Alternate Hypothesis

$$H_1: U > 0$$

At the 0.01 significance level, with 23 degrees of freedom, the critical t-statistic is 2.500

The data for \bar{X} yields a t-statistic = 8.0207

which is greater than the critical value and has a p-value less than 0.0001.

(p-value calculated using software available on <http://www.graphpad.com/quickcalcs/>)

Therefore at the 0.01 significance level we reject the null hypothesis, and conclude that this sample provides significant statistical evidence that in the population, the average difference of the Incongruent response time minus the Congruent response time is greater than zero.

Based on this sample, the 99% confidence interval estimate of U is (5.4824, 10.4476).

(Statistical analysis and visualizations were computed with LibreOffice Calc.)

6. Optional

It is likely that the participants' responses are slower of the Incongruent chart because they are processing both the color of the font and the name written in that color. Stroop's original research indicated that the name and color were being processed by two different parts of the brain. (Stroop, J.R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.*, 18:643-662, 1935. Referenced at <https://faculty.washington.edu/chudler/words.html#seffect>)

It could be possible to isolate the factor of the test subject's familiarity of the language printed in the Stroop chart. One possible test design would be to perform the Stroop experiment with the printed language being completely foreign to the test subject. If the subject were completely unfamiliar with the language printed, then one would expect the two responses, congruent and incongruent, to be close or nearly the same for each subject. The Stroop experiment has been conducted in such a fashion (with foreign language, and also with bi/multi-lingual readers) by numerous psychological researchers. For an example of such a variation of the Stroop experiment, please refer to *Stroop Effect Differences of Native and Non-Native Japanese Speakers*, Tiesling-Rusch and Dimond (<http://www.kon.org/urc/v11/rusch.html>).

Sources

<http://www.graphpad.com/quickcalcs/>

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

<http://www.kon.org/urc/v11/rusch.html>