Stroop Effect Experiment

by Fred W. Y. Toh

The following are the tasks/questions given in the instructions and my responses to them.

1. What is our independent variable? What is our dependent variable?

The independent variable in this experiment is the ink color of the words in the two test conditions. The dependent variable is the time taken for the participant in naming the ink colors in the two equally-sized lists.

2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

Null hypothesis, H_0:	The mean time taken in the congruent words condition, μ_C, is the same as the mean time taken in the incongruent words condition, μ_I.
Alternative hypothesis, H_A:	The mean time taken in the congruent words condition, μ_C, is shorter than the mean time taken in the incongruent words condition, μ_I.

Since the experiment depends on the same participant in both the congruent and incongruent words conditions, the paired-sample t-test is the appropriate statistical test. The assumptions taken are:

1. The sample size is less than 30.
2. The population standard deviation is unknown.
3. The distributions of the time taken for the two conditions are assumed Gaussian.

It is expected that the task in the incongruent words condition is tougher, and so we expect that it takes a longer time to complete compared to the congruent words condition task. Thus, a one-tailed test is the appropriate choice for the paired sample t-test we are about to conduct.

Now it's your chance to try out the Stroop task for yourself.

The following are the timings I achieved for the experiment:

time taken in congruent words condition, x_C = 17.28 s

time taken in incongruent words condition, x_I = 26.44 s

time difference, x_D = -9.16 s

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

The following are as calculated from the dataset. See workings attached for details.

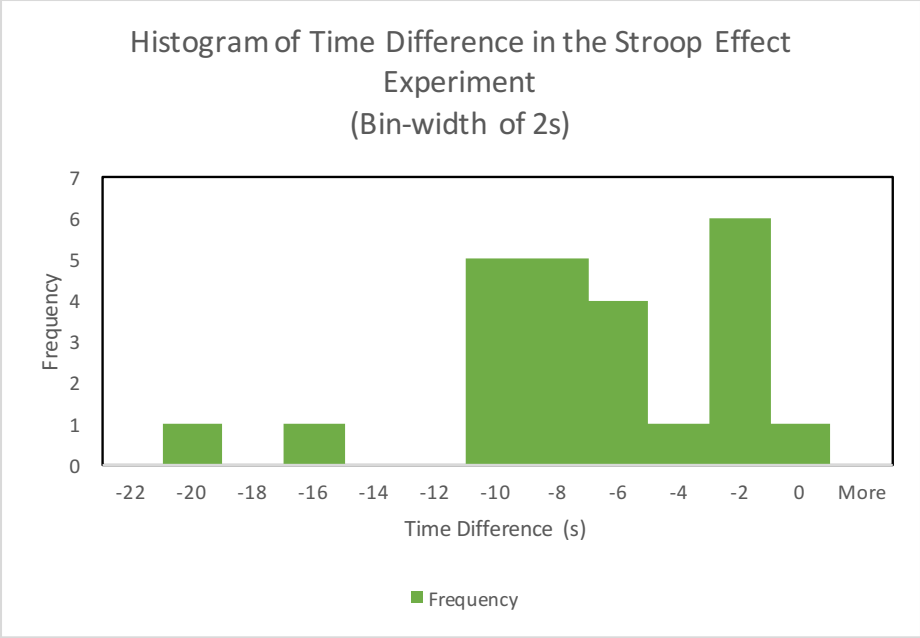
sample mean time taken in congruent words condition, $\bar{x}_C$ = 14.05 s

sample mean time taken in incongruent words condition, $\bar{x}_I$ = 22.02 s

mean time difference between the two conditions from the sample, $\bar{x}_D$ = -7.97 s

standard deviation of the differences from the sample, s_D = 4.87 s

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 histogram in the previous page shows the time differences of the participants in the dataset. We note that the sample size is small and there is no distinct “bell curve” shape in the distribution. There also seem to be a couple of extreme values.

5. Now, perform the statistical test and report your results. 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?

At level $\alpha = .05$,

Null hypothesis, H_0: $\mu_C - \mu_I \geq 0$

Alternative hypothesis, H_A: $\mu_C - \mu_I < 0$

Under t-statistics:

$$\text{t-score} = (\bar{X}_D - (\mu_C - \mu_I))/se, \qquad \text{for df} = n-1$$

where

$\bar{X}_D$ = variable for the mean time difference between the two conditions from the sample

se = standard error of the mean where $se = s_D / \sqrt{n}$

s_D = standard deviation of the differences from the sample

n = sample size

df = degrees of freedom

From the dataset, we compute the following:

$\bar{x}_D = -7.97$

$s_D = 4.87$

$n = 24$

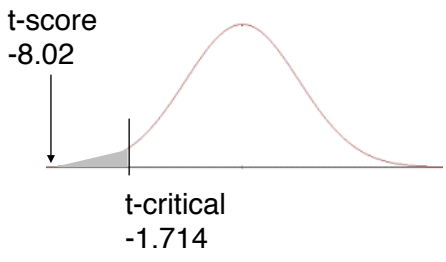
$se = 0.99$

t-score = -8.02

p-value < .0001

At level $\alpha = .05$ and df = 23,

$t\text{-critical}_{(23,0.05)} = -1.714$



Since the t-score of the sample distribution falls below the t-critical value ($p < .05$), we reject the null hypothesis at 95% confidence level in favor of the alternative that the mean time taken in the congruent words condition is shorter than that in the incongruent words condition. The outcome of this hypothesis test matches up with the choice of one-tailed test.

Computing at 90% confidence interval on the mean time difference of -7.97 s at the same α-level of .05, we have the following:

90% CI = (-9.67,-6.26)

My recorded time difference at -9.16 s falls within the 90% confidence interval of the sample mean.

6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect?

According to Eric H. Chudler, Ph.D. (See Ref. 3), there are two theories that may explain why it takes less time in the first part than the second part of the experiment. Note that in the first part, the participants may plainly read off the words since each color already matches the word even though the task is technically to name the color, not read the word. Thus, there is the Speed of Processing Theory, which explains that it is plainly faster to read words than to name colors. The second theory is the Selective Attention Theory, which explains that it is not necessarily true that it is faster to read words than to name colors but when the brain is faced with words and colors, it automatically prefers to process the words, even though the instruction is solely on the colors. And because of this conflict between the auto-instruction versus the external instruction, it causes the brain to be confused, and thus slows down the whole process.

For me, I wonder if there is a third component to the Stroop Effect that has not been explained for. In this case, we are presented with words that describe colors, which at the same time, we are tasked to deal with the colors, but not the words. It is evident that the association of the color words and the colors exists in the experiment. I note that the words are distracting even

as I am intentionally shifting my attention on the colors. Thus I think the word/color association is also interfering the brain processing in this particular experiment.

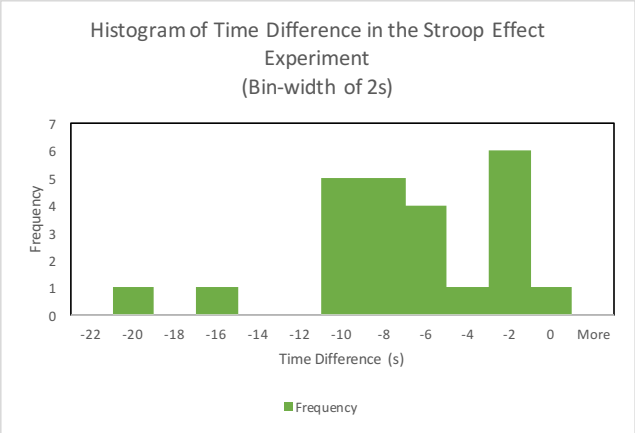
Chudler suggests modifying the experiments, such as turning the words upside down or rotating them 90 degrees. I believe in this way we may then be able to control for the association effect. I think another approach could be switching out the color words with other kinds of words, say random adjectives, to see if the Stroop Effect persists.

EXTERNAL REFERENCES:

1. Data set of recorded timings stroopdata.csv, provided by uDacity
2. Stroop Effect experiment https://faculty.washington.edu/chudler/java/ready.html
3. Neuroscience for Kids https://faculty.washington.edu/chudler/words.html

Congruent	Incongruent	Difference	Squared Deviation From Mean	
12.079	19.278	-7.199	0.586	
16.791	18.741	-1.950	36.178	
9.564	21.214	-11.650	13.581	
8.63	15.687	-7.057	0.824	
14.669	22.803	-8.134	0.029	
12.238	20.878	-8.640	0.456	
14.692	24.572	-9.880	3.668	
8.987	17.394	-8.407	0.196	
9.401	20.762	-11.361	11.534	
14.48	26.282	-11.802	14.724	
22.328	24.524	-2.196	33.279	
15.298	18.644	-3.346	21.333	
15.073	17.51	-2.437	30.556	
16.929	20.33	-3.401	20.828	
18.2	35.255	-17.055	82.632	
12.13	22.158	-10.028	4.257	
18.495	25.139	-6.644	1.744	
10.639	20.429	-9.790	3.331	
11.344	17.425	-6.081	3.549	
12.369	34.288	-21.919	194.720	
12.944	23.894	-10.950	8.911	
14.233	17.96	-3.727	17.959	
19.71	22.058	-2.348	31.548	
16.004	21.157	-5.153	7.906	
			544.33	SUM
14.051	22.016	-7.965		AVERAGE

Bin	*Frequency*
-22	0
-20	1
-18	0
-16	1
-14	0
-12	0
-10	5
-8	5
-6	4
-4	1
-2	6
0	1
More	0



DESCRIPTIVE STATISTICS:

median difference	-7.667	s
mean difference	-7.965	s
sample size, n	24	
sample std. dev. of the differences	4.865	s

INFERENCE STATISTICS:

standard error of the mean	0.993
t-score	-8.021
df	23
alpha	0.05
t-crit, (23,0.05)	-1.714
p-value of t-score	< 0.001

EFFECT SIZE MEASURE:

Cohen's d	-1.649
r^2	0.737

90% CONFIDENCE INTERVAL:

alpha	0.05	
mean	-7.965	s
standard error	0.993	s
t-crit, (23,0.05)	1.714	
Lower limit	-9.667	s
Upper limit	-6.263	s

OWN EXPERIMENT RESULTS:

Congruent	17.283	s
Incongruent	26.439	s
Time difference	-9.156	s
Conclusion:	Falls within the 90% confidence interval	