Stroop Effect Project

Introduction to Stroop Effect

In a Stroop task, participants are presented with a list of words, with each word displayed in a color of ink. The participant's task is to say out loud the color of the ink in which the word is printed.

The task has two conditions:a **congruent words** condition, and an **incongruent words** condition.

- In the *congruent words* condition, the words being displayed are color words whose names match the colors in which they are printed: for example **RED**,**BLUE**.
- In the *incongruent words* condition, the words displayed are color words whose names do not match the colors in which they are printed: for example **PURPLE**,**ORANGE** .

In each case, we measure *the time it takes to name the ink colors* in equally-sized lists.
Each participant will go through and record a time from each condition.

Questions For Investigation

Question 1: Identify variables in the experiment

Firstly, we need to identify the independent and dependent variables in the experiment.

Dependent variable: the time it takes to name the ink colors
Independent variable: type of words condition (congruent and incongruent)

Question 2a: Establish hypotheses

We try to understand if the mean difference of two conditions are statistically significant. With intuation:

Null hypothesis: For incongruent words, there is no difference on the time it takes to name the ink colors or the time it takes shorter than congruent words.

$$H_0 : \mu_i - \mu_c \leq 0$$

Alternative hypothesis: For incongruent words, the time it takes longer than congruent condition.

$$H_A : \mu_i - \mu_c > 0$$

μ_c is the population mean of the first condition "congurent words". μ_i is the population mean of the second condition "incongruent words". Our null hypothesis states that there is no difference between means or incongruent condition is less. If we reject null hypothesis, difference between means are statistically significant and incongruent condition is more(positive).

Question 2b: Establish a statistical test

We need to use t-test as opposed to z test because we do not know population parameters. We know parameters about samples with size=24.

Our type of study will be dependent samples t-test(paired t-test). Participants are same in each test and tested in different condition. Due to selecting hypothesis for one direction, we will use one-tailed t-test through positive side.

α-*level*= *.05* will be our selection.

Question 3: Report descriptive statistics

- Sample mean of congruent condition($\bar{x}_c$) : 14.051125
- Sample standart deviation of congruent condition (S_c) : 3.559357958
- Sample mean of incongruent condition ($\bar{x}_i$) : 22.01591667
- Sample standart deviation of congruent condition (S_i) : 4.797057122
- Mean of differences ($\bar{x}_d$) : $\bar{x}_i$ - $\bar{x}_c$ = 7.964791667
- Sample standart deviation of mean differences (S_d) : 4.86482691

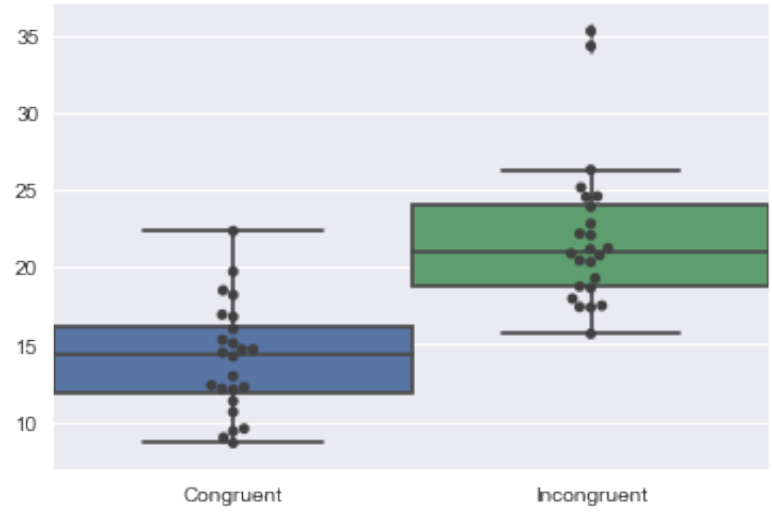
Part. No.	Congruent	Diff (Mean - Part.)	Square Diff
1	12.079	-1.972125	3.88927702
2	16.791	2.739875	7.50691502
3	9.564	-4.487125	20.1342908
4	8.63	-5.421125	29.3885963
5	14.669	0.617875	0.38176952
6	12.238	-1.813125	3.28742227
7	14.692	0.640875	0.41072077
8	8.987	-5.064125	25.645362
9	9.401	-4.650125	21.6236625
10	14.48	0.428875	0.18393377
11	22.328	8.276875	68.5066598
12	15.298	1.246875	1.55469727
13	15.073	1.021875	1.04422852
14	16.929	2.877875	8.28216452
15	18.2	4.148875	17.2131638
16	12.13	-1.921125	3.69072127
17	18.495	4.443875	19.748025
18	10.639	-3.412125	11.642597
19	11.344	-2.707125	7.32852577
20	12.369	-1.682125	2.82954452
21	12.944	-1.107125	1.22572577
22	14.233	0.181875	0.03307852
23	19.71	5.658875	32.0228663
24	16.004	1.952875	3.81372077
Mean of cong.	14.051125	SS	291.387669
		n-1	23
		Variance	12.6690291
		Congruent Sample SD	3.55935796

Part. No.	Incongruent	Diff (Mean - Part.)	Square Diff
1	19.278	-2.737916667	7.49618767
2	18.741	-3.274916667	10.7250792
3	21.214	-0.801916667	0.64307034
4	15.687	-6.328916667	40.0551862
5	22.803	0.787083333	0.61950017
6	20.878	-1.137916667	1.29485434
7	24.572	2.556083333	6.53356201
8	17.394	-4.621916667	21.3621137
9	20.762	-1.253916667	1.57230701
10	26.282	4.266083333	18.199467
11	24.524	2.508083333	6.29048201
12	18.644	-3.371916667	11.369822
13	17.51	-4.505916667	20.303285
14	20.33	-1.685916667	2.84231501
15	35.255	13.23908333	175.273328
16	22.158	0.142083333	0.02018767
17	25.139	3.123083333	9.75364951
18	20.429	-1.586916667	2.51830451
19	17.425	-4.590916667	21.0765158
20	34.288	12.27208333	150.604029
21	23.894	1.878083333	3.52719701
22	17.96	-4.055916667	16.45046
23	22.058	0.042083333	0.00177101
24	21.157	-0.858916667	0.73773784
Mean of incong.	22.01591667	SS	529.270412
		n-1	23
		Variance	23.011757
		Incongruent Sample SD	4.79705712

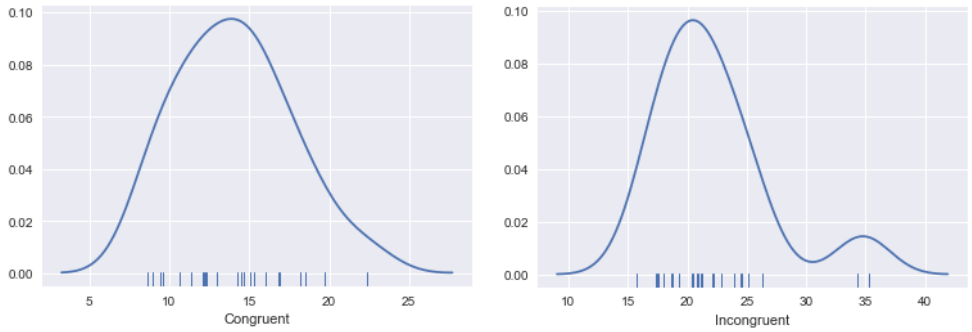
Part. No.	(Incong.-Cong.)	Diff(Mean-(I-C))	Square Diff
1	7.199	-0.765791667	0.586436877
2	1.95	-6.014791667	36.17771879
3	11.65	3.685208333	13.58076046
4	7.057	-0.907791667	0.82408571
5	8.134	0.169208333	0.02863146
6	8.64	0.675208333	0.455906293
7	9.88	1.915208333	3.66802296
8	8.407	0.442208333	0.19554821
9	11.361	3.396208333	11.53423104
10	11.802	3.837208333	14.72416779
11	2.196	-5.768791667	33.27895729
12	3.346	-4.618791667	21.33323646
13	2.437	-5.527791667	30.55648071
14	3.401	-4.563791667	20.82819438
15	17.055	9.090208333	82.63188754
16	10.028	2.063208333	4.256828627
17	6.644	-1.320791667	1.744490627
18	9.79	1.825208333	3.33138546
19	6.081	-1.883791667	3.548671043
20	21.919	13.95420833	194.7199302
21	10.95	2.985208333	8.911468793
22	3.727	-4.237791667	17.95887821
23	2.348	-5.616791667	31.54834863
24	5.153	-2.811791667	7.906172377
Mean of Diff.	7.964791667	SS	544.33044
		n-1	23
		Variance	23.66654087
		Mean Diff. Sample SD	4.86482691

Question 4: Plot the data

Boxplot for participant data of two conditions



Kernel density plot for participant data of two conditions



We can observe mean and quantiles placements in the boxplot. Incongruent sample is placed with higher values than congruent sample. Distributions are likely normal distribution except outliers at incongruent sample.

Question 5: Perform the statistical test and interpret your results

- $\bar{x_d}$: 7.964791667
- S_d : 4.86482691
- **n**= 24
- **df**=n-1=23

According to α-*level*= *.05* for *one-tail* hypothesis and *df=23*, we can reference from t-table:

$$t_{\text{critical}} = 1.714$$

Calculate t-statistic with $\bar{x_d}$, S_d and n:

$$t = \frac{\bar{x_d}}{S_d / \sqrt{n}} == 8.020706944$$

Interpretion

t-statistic is greater than $t_{critical}$ and inside of critical region, so **p-value < .05** . Calculation of p-value with df and t-value in graphpad website results p value is smaller than 0.0001.

As a result, we will reject null hypothesis($ H_0 $). This means that testing with incongruent words, the time it takes longer than congruent condition. This extra time is statistically significant. Result is match up our intuition at the beginning.

Calculate r^2:

$$r^2 = \frac{t^2}{t^2 + df} = 0.73663608$$

This means 73.66 of variability in time of strop test is due to word condition(congurent - incongruent).

Question 6: Digging deeper and extending the investigation

Possible explanation, people tend to comprehense information more quickly if they are compatible with senses. If there is inconsistency and mismatch, people may need to check to avoid mistake and this may results to consume time. In this experiment, colorful words with color names are examples of distraction. There should be another expriment with colorful words but not include "color names" and it could be also compared effects.

Similar test may prepared for direction and words by placing words inside of a big direction arrow.For example, arraw shows "UP" direction while "DOWN" writes inside of it.

Resources:

- t-table : https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg
- Data: https://drive.google.com/file/d/0B9Yf01UaIbUgQXpYb2NhZ29yX1U/view
- p-value calculation : https://www.graphpad.com/quickcalcs/pValue2/
- Stroop task: https://faculty.washington.edu/chudler/java/ready.html