Statistics: The Science of Decisions
Project Instructions

Background Information

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:

1) a congruent words condition, and an 2) 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

As a general note, be sure to keep a record of any resources that you use or refer to in the creation of your project. You will need to report your sources as part of the project submission.

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

Dependent: the time it takes to name the ink colors (in equally sized lists, for the conditions, for the participants).

Independent: the two conditions. The condition "Congruent" (word colour match) and the condition "Incongruent" (word colour mismatch).

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

We have here an experiment in which the same users take the test twice, first with the Congruent and then with the Incongruent set. Therefore, we have related groups (the same subjects are present in both tests), so we could look for differences between means when the same subjects are measured on the same dependent variable under two different conditions.

The statistical test that fits here perfectly is a dependent t-test for two conditions for paired samples.

With the information provided, the most reasonable hypothesis would be to check if there are significate differences between the two conditions.

We will perform the statistical set assuming that the introduction of the Incongruent set will produce a variability of the results, and not an increase or decrease in the time of response.

Null Hypothesis $\mathbf{H_o}$: the reading time will be equal for the two different conditions.

Alternative Hypothesis $\mathbf{H_1}$: the reading time will be not equal for the two different conditions.

Being X_{cong} the mean of the time spent by the participants on naming the colors at the congruent condition and X_{incong} the mean of the time spent by the participants on naming the colors at the incongruent condition, we could express the Null and Alternative Hypothesis as follows:

$\mathbf{H_o}$: $\mu_{cong} = \mu_{incong}$ or $\mu_{cong} - \mu_{incong} = 0$ or $\mu_D = 0$ → equal time

$\mathbf{H_1}$: $\mu_{cong} \neq \mu_{incong}$ or $\mu_{cong} - \mu_{incong} \neq 0$ or $\mu_D \neq 0$ → different time

two-tailed t-test

But we could have tested also:

H_o: $\mu_{cong} \geq \mu_{incong}$ or $\mu_{cong} - \mu_{incong} \geq 0$ → no effect, or less time

H_1: $\mu_{cong} < \mu_{incong}$ or $\mu_{cong} - \mu_{incong} < 0$ → effect, more time

one-tailed t-test (+)

That means that we expect an increase of the response time in the Incongruent condition vs. the Congruent condition.

---

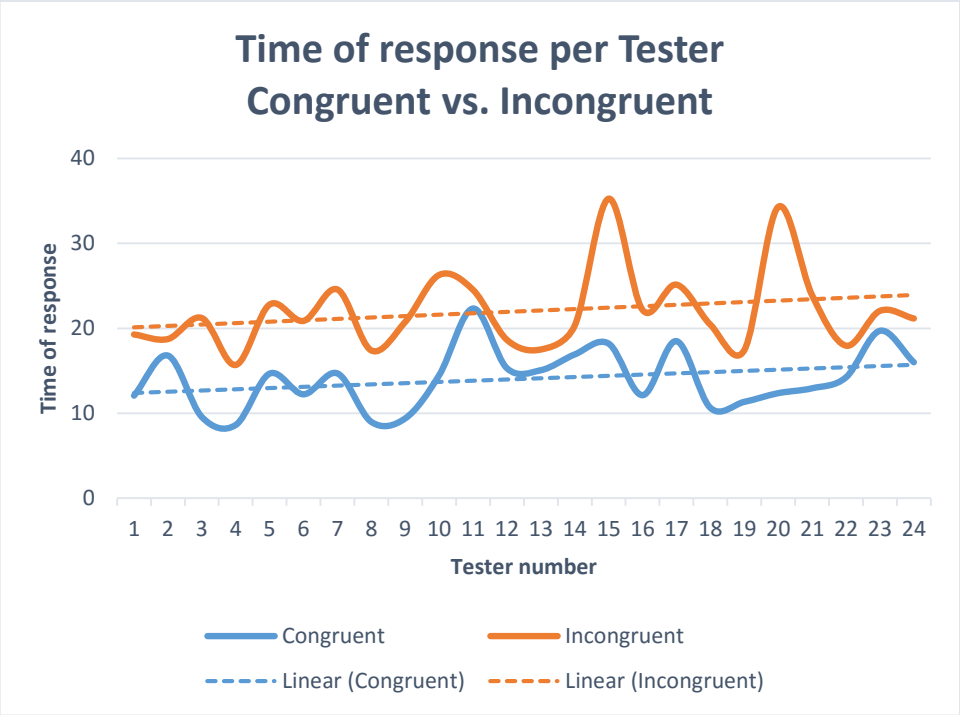
3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.

Central tendency: median, mode, mean

Measure of variability: standard deviation, variance and range

	Congruent	*Incongruent*
Mean	14.051125	22.01591667
Standard Error	0.726550901	0.979195185
Median	14.3565	21.0175
Mode	12, 15 (if no decimals)	21 (if no decimals)
Standard Deviation Sample	3.559357958	4.797057122
Sample Variance	12.66902907	23.01175704
Kurtosis	-0.205224823	2.688900198
Skewness	0.416899874	1.547590026
Range	13.698	19.568
Minimum	8.63	15.687
Maximum	22.328	35.255
Sum	337.227	528.382
Count	24	24
Confidence Level(95.0%)	1.50298505	2.025619571

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 graph above shows two visualizations in one: the solid lines show the time of response for all participants (24), being the “Congruent” set blue and the “Incongruent” set orange.

The dotted line shows the trendline for “Congruent” set (blue) and with the “Incongruent” set (orange).

We can clearly see that the time of response in the “Congruent” case is superior to the time of response in the “Incongruent” case, orange dotted line above the blue dotted line.

On average, based on the graph, we could affirm that it takes more time to speak aloud the Incongruent set of words that the Congruent one.

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?*

As explained in the answer to question 2, the statistical test to be performed is a dependent t-test for two conditions for paired samples.

We will test it with a 95% confidence level ($\alpha = 0.05$) two-tailed, df 23, $t_{critical}$ = +- 2.069 (otherwise, one-tailed + 1.714).

Our calculated $\mathbf{t_{stat}}$ is **8.03** (calc. "by hand", the software reported 8.02).

Conclusion: whether we perform a one or a two-tailed test, the value of the $t_{stat} > t_{critical}$ (by far) and as a consequence we reject the null Hypothesis and assume the alternative Hypothesis. Also, $p < 0.05$ in both cases (see table below).

In other words, there is statistical evidence that the introduction of the Incongruent set will produce longer response time.

Now, let us check a couple more statistics:

Confidence interval: (**5.92 – 10.02**). On average, the users will take between 5.92 and 10.02 units of time more to name the Incongruent words than the Congruent ones.

$\mathbf{R^2}$: (0.7373 or **74%**). The changes of conditions explain the 74% of the variability (relatively high).

More figures:

S = sqrt [$(\Sigma(d_i - d)^2 / (n - 1)$] = sqrt [(544.33 / (24- 1)] = **4.86** **----** (d = mean of the differences, 7.96)

SEM= s / sqrt(n) = 4.86 / sqrt(24) = **0.992**

$\mathbf{T_{stat}}$ = [$(x_1 - x_2) - D$] / SE = (d - D)/ SE = (7.96-0) / 0.992 = **8.024**

Margin of error: +- **2.052**

Excel output:

	Variable 1	*Variable 2*
Mean	22.01591667	14.051125
Variance	23.01175704	12.66902907
Observations	24	24
Pearson Correlation	0.351819527	
Hypothesized Mean Difference	0	
df	23	
t Stat	8.020706944	
P(T<=t) one-tail	2.0515E-08	
t Critical one-tail	1.713871528	
P(T<=t) two-tail	4.103E-08	
t Critical two-tail	2.06865761	

SPSS output:

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 I - C	7.964792	4.864827	.993029	5.910555	10.019028	8.021	23	.000

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? Some research about the problem will be helpful for thinking about these two questions!*

The responsible of the effects observed is the way the human brain is "hard wired". It, of course, finds an optimal solution but it just takes more time in the majority of the individuals.

Another influence could be the way we are educated from childhood to read words and not colors (just imagine and alphabet of colors where the letters are just shapes)….

Alternative tests could be made with different background colors (http://www.swarthmore.edu/SocSci/fdurgin1/ReverseStroop/PBRStroop.html).

In the Wikipedia, there are examples of the modifications made to the Stroop test including other sensory modalities and variables modifications (for example, warped words, emotional, spatial, numerical, reverse…etc modifications of the conditions) to study for example the effects of bilingualism, the effect of the emotions on interference and so on (https://en.wikipedia.org/wiki/Stroop_effect#Variations).