

Test a perceptual phenomenon: the « Stroop effect »

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: 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.

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

Independent variable is the word condition (congruent or incongruent).

Dependent variable is the time it takes to name the ink colors of the words displayed.

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

In order to compare results of the « Stroop effect » test and figure out whether test conditions has a significant impact on the time to carry out the test, it is proposed to carry out a statistical test.

Type:

Results for the tests (congruent words and incongruent words conditions) will be limited in size and not representative of the entire population. Therefore the statistical test to compare the mean results of each test should be a **two independent sample t-test**.

Direction:

We don't know if a test condition has a favorable or unfavorable effect on the time to carry out the test. Therefore we chose to perform a **two tailed test**.

Correction for unequal variance:

It is thought that test results will have unequal variance, which has to be taken into account in the computation of the standard error. **Welch's t-test** is used to take into account that feature.

Null and alternative hypothesis:

The null hypothesis is that test condition (congruent or incongruent) has no impact on the test results (time to complete the test).

$$H_0 : \mu_{congruent} - \mu_{incongruent} = 0$$

The alternative hypothesis is that the test condition has an impact on the test results

$$H_A : \mu_{congruent} - \mu_{incongruent} \neq 0$$

In above equations

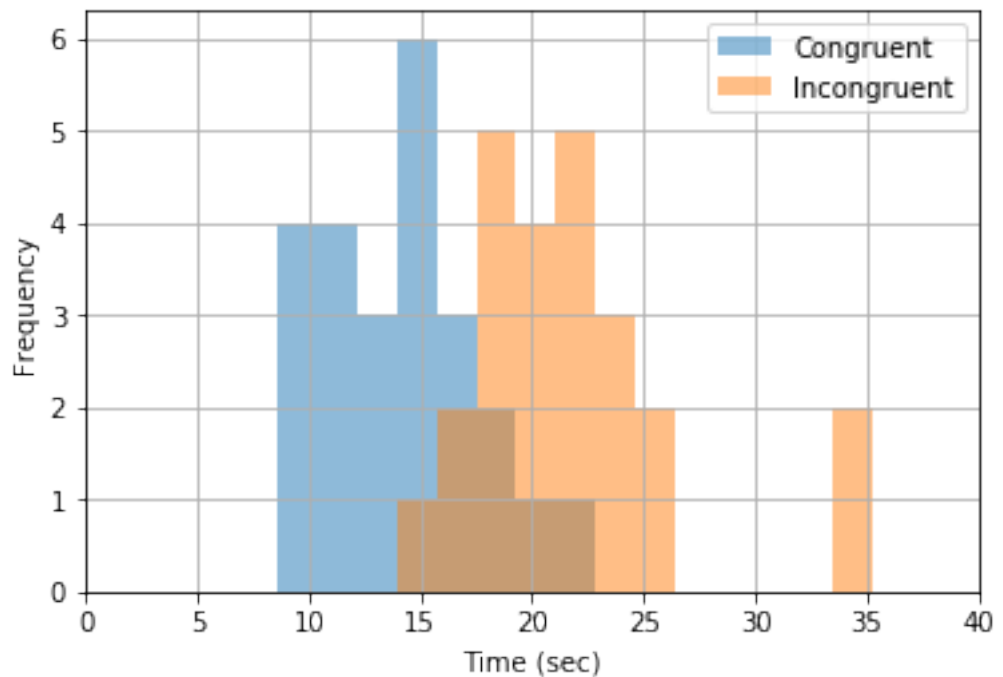
- μ stands for sample mean
- H_0 stands for the null hypothesis
- H_A stands for the alternative hypothesis

Threshold:

In order to reject the null hypothesis, a **significance level is set at 5%**. That means, the null hypothesis will be rejected if the probability that there is a difference between the test results is higher than 95%.

3 & 4: Descriptive statistics and visualisation of the distribution of the sample data

Following figure provides a description of the test results for congruent words and incongruent words conditions.



	Congruent	Incongruent
Number of samples	24.0	24.0
Min (sec)	8.63	15.69
Max (sec)	22.33	35.26
Mean (sec)	14.05	22.02
Median (sec)	14.36	21.02
Sample std (sec)	3.56	4.8

The visualisation of the distribution of the sample data (histogram), shows that both test results show a normal distribution with some overlap.

The descriptive statistics analysis shows that:

- in average it takes 8 more seconds to carry out the incongruent test compared to the congruent test
- sample standard deviation for both tests is around 4 seconds

An inferential statistical analysis is required to conclude whether results of both tests are significantly different.

5. Statistical test

A two tailed, two independent sample t-test with unequal variance is perform on the dataset described above.

Python code output:

```
alpha: 0.05
degree of freedom: 46
t_statistic: -6.532
t_critical: -2.013
hypothesis testing result: reject H0
r2: 48%
margin of error: 2.45
congruent mean: 14.05
incongruent mean: 22.02
confidence interval for congruent mean: (11.60,16.51)
```

Result of the two independent sample t-test is

$t(46) = -6.532$, $p < .05$, two tailed

We reject the null hypothesis and conclude there is a significant difference between the two sample results.

The confidence interval at 95% for the congruent test is (11.60 sec to 16.51 sec) and the incongruent test takes a longer time (22.02 sec in average).

48% of the difference between the two test results can be explained by the test conditions.

Results are in line with my expectation and with my experience of the test. In fact it took me longer to tell the colors when they were not matching the words.

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!

My understanding of the phenomena is that human brain have different perception capabilities and when those are in conflict it takes more time to process. In the present case, the two perception capabilities are word reading and color identification. When both informations are aligned (congruent words condition), brain processing time is quick while when informations are different (incongruent words condition) it takes more time.

I've made some tests at home with my family and it did not work as well due to the fact the words were written in English. When I used the French version of the test (see figure below), the time difference between the congruent and incongruent conditions was much more important. I believe it is due to the fact that my daughters (9 and 7 years old) do not read English yet while they just learnt to read French. I would like to do the same test with my younger son (4 years old). My intuition is that there should be no difference between the English and the French test as well as between the congruent and the incongruent conditions because he does not know how to read yet. Of course a proper statistical test should be set up if I wanted to prove it.

vert bleu jaune rouge vert
bleu rouge jaune vert bleu
rouge jaune vert rouge bleu
jaune vert bleu jaune rouge
jaune bleu rouge vert jaune
rouge jaune vert rouge bleu
vert jaune rouge bleu vert
jaune rouge bleu bleu jaune
rouge jaune vert rouge bleu
vert bleu rouge jaune vert

The website from the faculty of Washington (<http://faculty.washington.edu/chudler/words.html>) proposes other « Stroop effect » test with direction, numbers and animals instead of colors. I've tested and found them easier than the color test. It is probably due to the fact that the brain capabilities in conflict (direction reading vs positioning, digit reading vs digits counting, word reading vs animal recognition) are of more even strength.