Project “Test a perceptual phenomenon”

Report by Jörg Strebel, 26.01.2018

Question 1: Identify variables in the experiment

The dependent variable is the time measured to name the ink colors in equally-sized lists.
The independent variable is the condition, whether the ink color and the word meaning are congruent or incongruent.

Question 2a: Establish hypotheses

The null hypothesis H_0 is that there is no time difference between the congruent words condition and the incongruent words condition. The alternative hypothesis H_a would assume that any person can more quickly process the congruent words condition and thus the time difference would be negative.

Null → H_0: $\mu_c - \mu_i >= 0$
Alternative → H_a: $\mu_c - \mu_i < 0$

μ_c is the population mean for the congruent words condition, μ_i is the population mean for the incongruent words condition.

Question 2b: Establish a statistical test

This is a within-sample test design with two conditions measured on the same participant. So, a paired-sample dependent t-test is the right test here. It tests whether two sample means are different. In this case, I use a one-tailed test in the negative direction.

The statistical assumptions of this test are as follows (c.f. [1]):

- participants were randomly sampled from the population.
- Two normally distributed populations for the two conditions; the difference between the paired values has to be approximately normally distributed.
- identical variance in the two populations

The assumption of normality can be checked later using a histogram (or a Q-Q-plot).

Reasons for a t-test (instead of a z-test for example):

- I am trying to compare the sample means of two dependent samples and test whether they are from the same population or from two different populations. If I wanted to use a z-test, I would

have to compare my sample mean with the population mean and I would need to know both the population mean and the population standard deviation (which both are not available in this case).
- As I don't know the population standard deviation, I have to estimate the sample standard deviation from the sample data itself, which necessitates the use of a t-distributed test statistic.
- As I have only a limited number of observations, the above mentioned test statistic only has a limited number of degrees of freedom, which has to be considered when choosing the statistical test.

Question 3: Report descriptive statistics

In the below table, stroop_df_long is the name of the R data frame, and Condition is the variable containing the name of the condition.

```
stroop_df_long$Condition: Congruent
   Min.     1st Qu.     Median      Mean        3rd Qu.     Max.
   8.63     11.90       14.36       14.05       16.20       22.33
------------------------------------------------------------------------------
stroop_df_long$Condition: Incongruent
   Min.     1st Qu.     Median      Mean        3rd Qu.     Max.
  15.69     18.72       21.02       22.02       24.05       35.26
```

	Congruent	**Incongruent**
Number of observations	24	24
Degrees of freedom	23	23
Sample Standard Deviation	3.559	4.797

The two groups show different values. The sample with the congruent condition shows overall smaller mean and median values and even a smaller sampler standard deviation.

Question 4: Plot the data

Two visualizations have been created (see below).

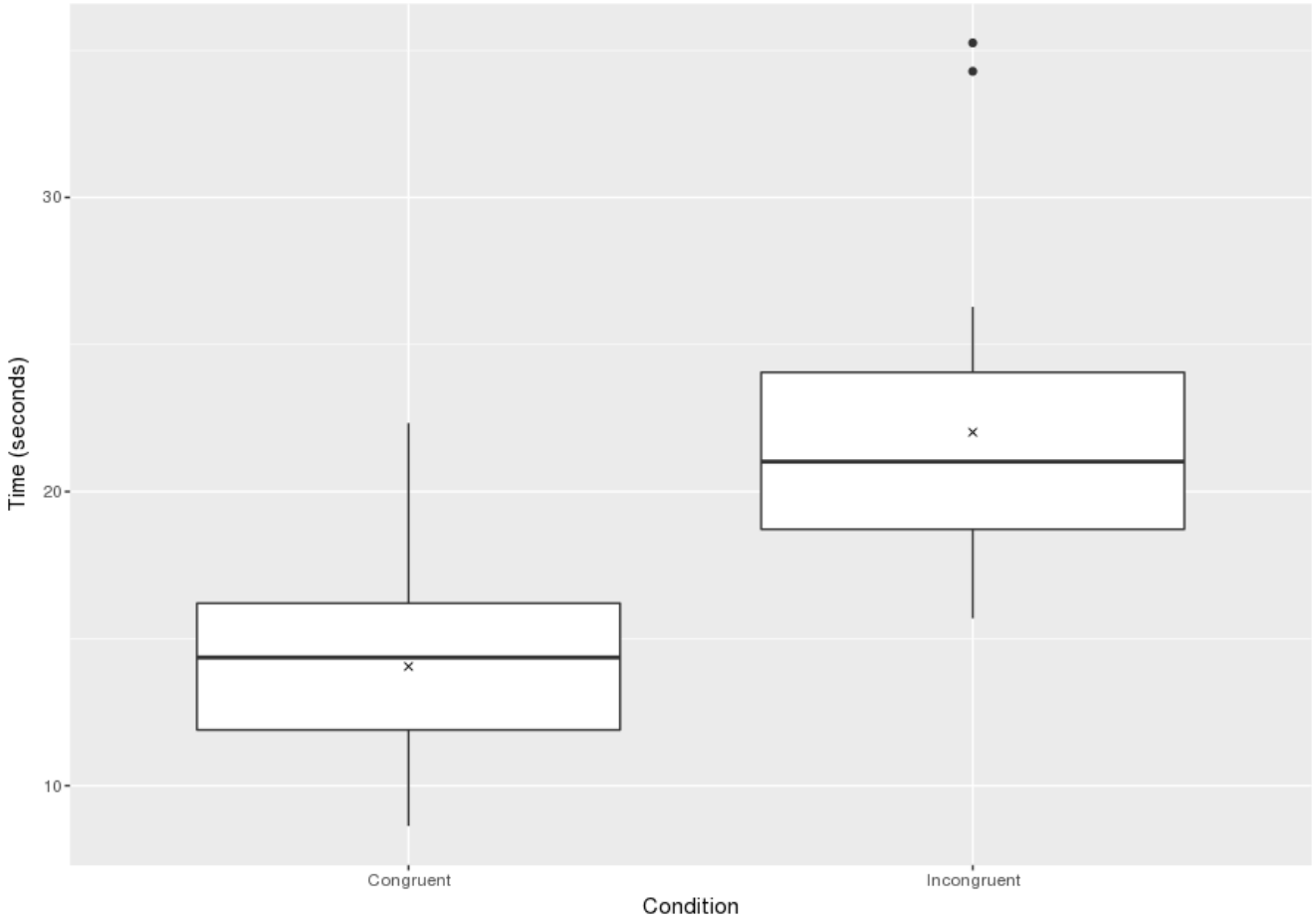

Illustration 1: Box plot

Illustration 1 shows a box plot of the data. The mean value for each group is marked as a cross in column. You can clearly see the difference in the data. There are two outliers in the incongruent sample; maybe there was an error of measurement, as these values look wrong.

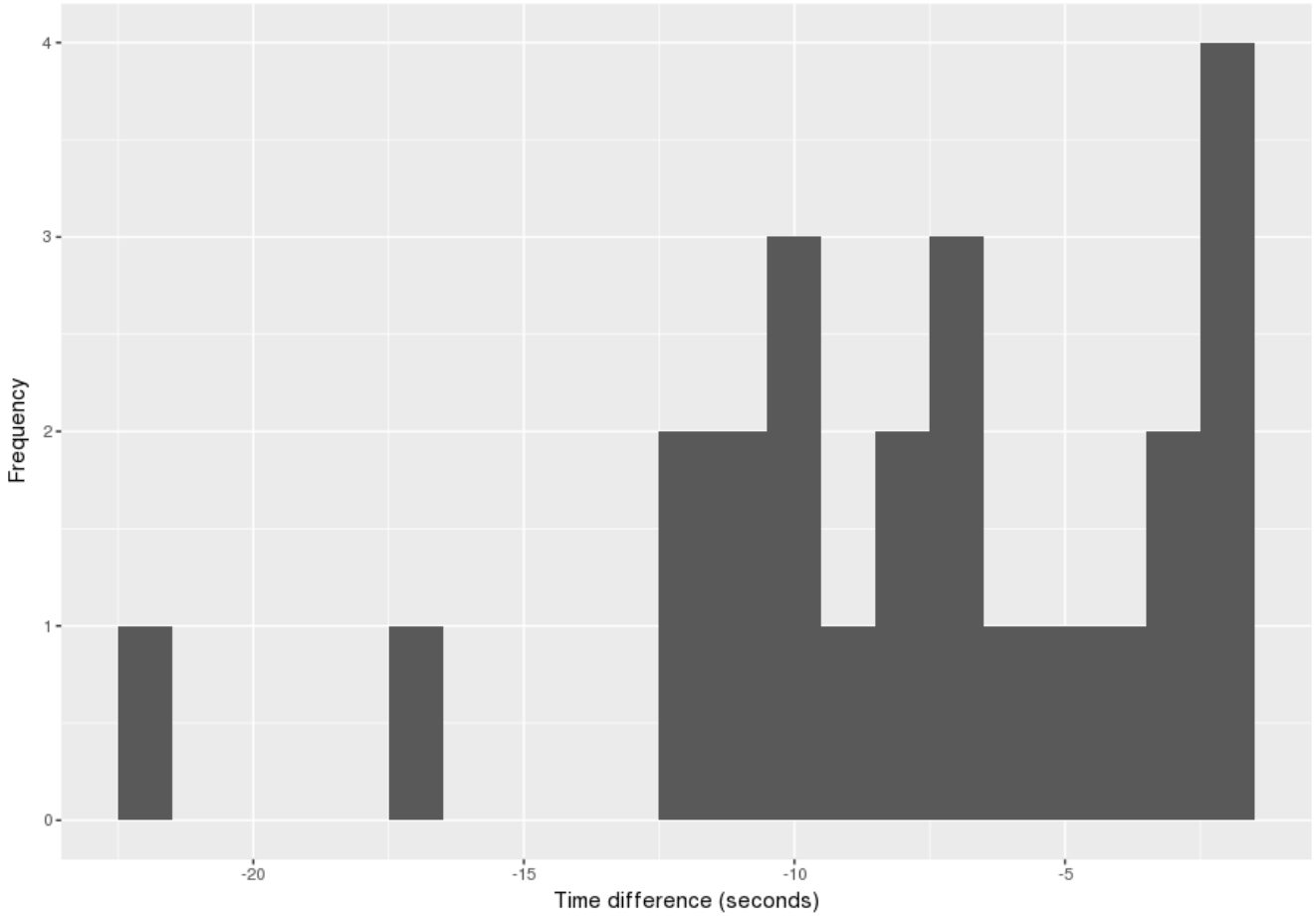

Illustration 2: Histogram

Illustration 2 shows a histogram of the differences between the timing values of the congruent and the incongruent sample. A binwidth of one second was used. It is not normally distributed, which technically prohibits the use of a t-test.

Question 5: Perform the statistical test and interpret your results

Significance level α = 0.05, Confidence level 95%, one-tailed
Critical value of t-statistic: -1.714 (23 degrees of freedom)

Sample standard distribution of the mean differences: 4.865
Standard error of the mean: 0.993

```
data:  stroop_df$Congruent and stroop_df$Incongruent
test-statistic = -8.0207,
df = 23,
p-value = 2.052e-08

alternative hypothesis: true difference in means is less than 0
```

```
95 percent confidence interval:      (-Inf, -6.262868)

sample estimates:
mean of the differences: -7.964792
```

As the test-statistic is clearly in the critical region, I would reject the null hypothesis; the two groups don't seem to come from the same population as their means are statistically different. As the very low p-value suggests, this result is very unlikely due to chance / sampling error.
It seems the congruent task is mentally easier for the brain than the incongruent task. This result matches with my expectation in two ways: first, the box plot already hinted at this result. Second, it takes focus to tell apart the ink color from the word meaning; if both are the same, less focus is needed, which should be faster.

Question 6: Digging deeper and extending the investigation

The web page [2] is a good resource for hypotheses regarding the reasons for the observed effect and also lists variations of the original experiment, that may produce similar effects. According to this web site, there are two theories that may explain the Stroop effect:

- Speed of Processing Theory: the interference occurs because words are read faster than colors are named.
- Selective Attention Theory: the interference occurs because naming colors requires more attention than reading words.

The above mentioned interference means the problem that the brain has with the two types of information (color vs. word meaning).

But there might be other theories that can provide a broader vision of the phenomenon and its possible reasons:

- Automation of reading theory/Automaticity Hypothesis [3]: Learning to read is a complex task involving considerable amount of practice across different stages of competency. The reason for the Stroop effect could be attributed to the fact, that the brain is trained for automatic reading; but in the Stroop experiment, there is no reading task, so the automaticity gets in the way of naming the colors.
- Bottleneck theory [4]: a class of theories of attention according to which all incoming information is subjected to some level of processing before a portion of it is selected for attention. It could be hypothesized that the color and the word meaning have to pass simultaneously through a common mental bottleneck which limits information perception.
- Parallel distributed processing theory [5]: The PDP model holds that the cognitive processes can be explained by activation flowing through networks that link together nodes. Every new event changes the strength of connections among relevant units by altering the connection weights. Consequently, you are likely to respond differently the next time you experience a similar event, as the neural network in the brain was already trained on the event the first time. But in the case of the Stroop effect, there are two conflicting stimuli, that could be hypothesized to conflict with each other.

Some related experiments might serve to elucidate the underlying causes for this phenomenon.
One such experiment controls for the word meaning. You could assume that the puzzle would be easier for a very young child than for older children or adults. The child should know the color names, but should not be able to read the words. If there is an interference in adults, it should disappear in children, as they cannot understand the word meanings [2].
Another class of experiments would try to subvert the reading process, so that the brain can more exclusively focus on the color. The test setup could be changed such that the words are turned upside down or rotated by 90 degrees, or that only half of the word or the first and last letter of each word is colored.
To better understand under what conditions the above mentioned interference becomes stronger or weaker, the test setup could be change such that non-color words such as "dog" or "house.” or nonsense words such as "kiw" or "thoz." are used. Also, you could compare long words to short words, or emotional words such as "sad" or "happy" or "depressed" or "angry."

References

[1] http://www.ncl.ac.uk/itservice/dataanalysis/simpletests/ttests/pairedsamplesttestdependentsamplesttest/
[2] https://faculty.washington.edu/chudler/words.html
[3] http://www.teach-nology.com/themes/lang_arts/reading/auto.html
[4] http://www.oxfordreference.com/view/10.1093/oi/authority.20110803095520271
[5] http://www.massey.ac.nz/~wwpapajl/evolution/assign2/YD/parallel.htm