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

The independent variable is the type of test, i.e. Congruent word or Incongruent word.
The dependant variable is the time that the participants took in the tests to recognize the color/word.

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

Since the samples are paired, i.e for each individual sample. there is a corresponding unique pair in the other sample we can use the following set of hypotheses:

H0: The null hypotheses that states that the intervention (congruent vs incongruent words) does **not have any effect** on the population's time to recognize the word.
$\mu i - \mu c = 0$ where μi is the samples for incongruent words and μc is the sample for congruent words

H1: The alternate hypotheses that states that **there is a significant difference** (+ve or -ve) in the average response time after the intervention.
$\mu i - \mu c \neq 0$

The suitable type of test for this kind of data would be a **dependent t-test**.
The *subjects in the test are same* but under different conditions. It also means that the subjects might have practice doing the second test based on their experience in the first test, or in other words, *the conditions are dependant*.
The samples are in pairs, meaning there is a sample for the same subject pre and post intervention.

hence we use **dependent/pair t-test** and **not independent t-test**

The reasons we cannot perform a z-test are:

- The data does not include information about the entire population
- When population size < 30, we cannot safely assume the sample standard deviation is a correct approximation of population standard deviation.

What we are trying to infer from this statistic is if congruent or incongruent word-color has any effect on the response time of the user identifying them. We are trying to infer this about the population using the sample data that we have collected via the experiment.

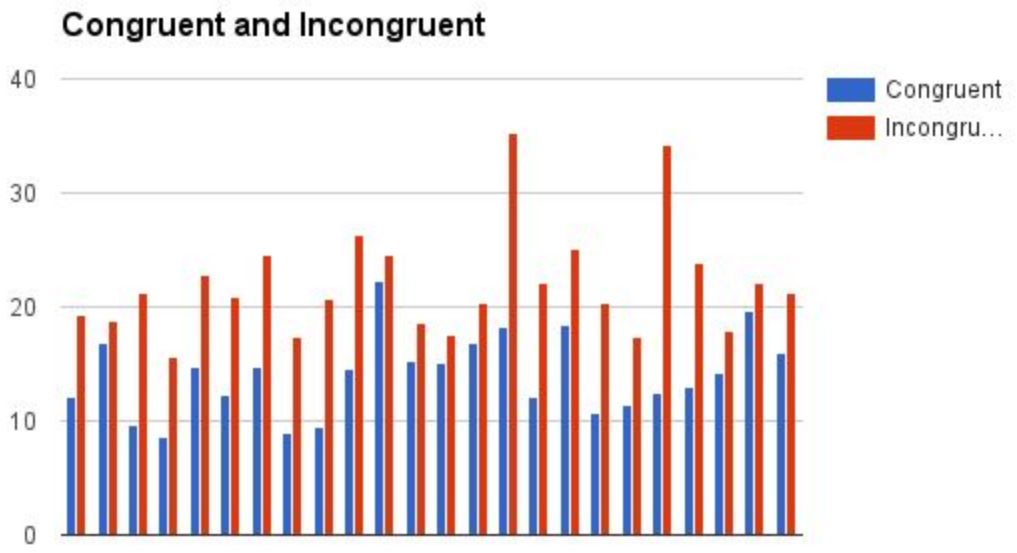
Since we are trying to hypothesize that there is a significant difference in the response time, either in positive or negative way, we need to use a **two-tailed test.**

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

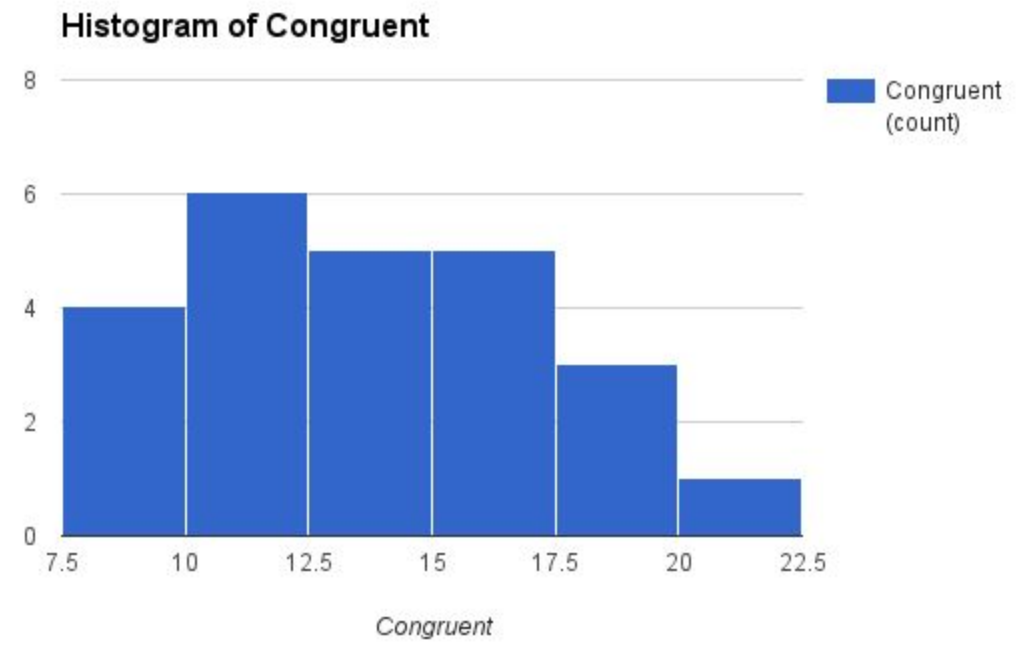
Statistic	Congruent Word	Incongruent Word
Mean	14.3565	21.0175
Standard Deviation	3.559357958	4.797057122
Variance	12.66902907	23.01175704
Median	14.3565	21.0175

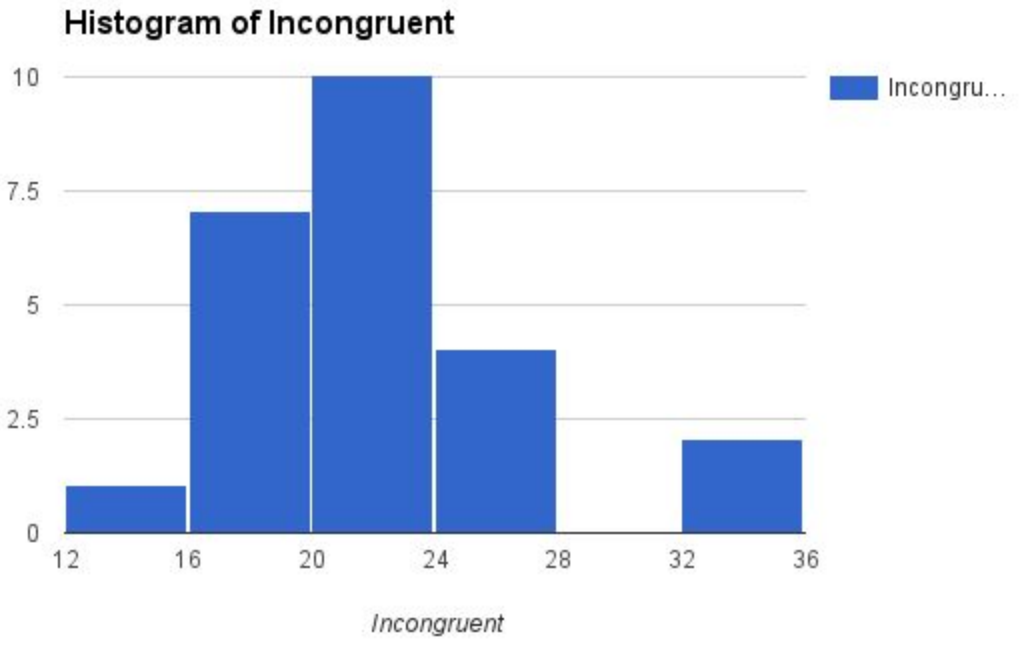
The mode cannot be determined for the dataset since none of the values repeat.

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 above visualization shows the variation in time between congruent and incongruent times. It clearly indicates that in all the cases, the incongruent time is higher than congruent time. This indicates that when names of the color are written in colors other than what the word says, it will most likely take longer for people to recognize it correctly.





The above two histograms indicate the time data. We can see the most between 10-12.5 was the most common congruent time and 20-24 was the most common incongruent time.
We can also infer from the histogram that 20-22.5 was the least common congruent time and 12-16 was the least common incongruent time. The interesting thing about this is that it can be

inferred that most people took at least 16 sec for incongruent test and very few people took more than 20 secs for congruent test.

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?**

Degree of freedom **df = n-1 = 24-1 = 23**
let **p-critical = 0.05**
from the table (https://people.richland.edu/james/lecture/m170/tbl-t.html) , for p-critical value of 0.05 and two-tailed test, **t-critical is 2.069**

so a value of ±2.069 for t would mean that we can reject the null hypotheses that the intervention does not have any significant change in the response time.

The value of t >= 2.069 or t<=-2.069 indicates that the difference in pre and post intervention is significant enough and that the intervention causes the response time to change.

Data:

Congruent	Incongruent	Difference
12.079	19.278	-7.199
16.791	18.741	-1.95
9.564	21.214	-11.65
8.63	15.687	-7.057
14.669	22.803	-8.134
12.238	20.878	-8.64
14.692	24.572	-9.88
8.987	17.394	-8.407
9.401	20.762	-11.361
14.48	26.282	-11.802
22.328	24.524	-2.196
15.298	18.644	-3.346
15.073	17.51	-2.437
16.929	20.33	-3.401
18.2	35.255	-17.055
12.13	22.158	-10.028

18.495	25.139	-6.644
10.639	20.429	-9.79
11.344	17.425	-6.081
12.369	34.288	-21.919
12.944	23.894	-10.95
14.233	17.96	-3.727
19.71	22.058	-2.348
16.004	21.157	-5.153

mean of difference = -7.964
Standard Deviation = 4.865
n = 24
Standard Error = SD/sqrt(n) = 4.865/sqrt(24) = 0.994

t = mean of difference / standard error = -8.012

since t < -2.069, **we can reject the null hypotheses with 95% confidence** (for df=23, p-critical = 0.05 and two-tailed test).

The results match up to the expectation I had based on the preliminary visualization and looking at the data.

Reference:

http://www.statisticslectures.com/topics/dependentsamplest/#video