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?

- Our independent variable is the list of words and colors (both congruent and incongruent)
- Our dependent variable is the time, in seconds, it takes for a participant to complete the Stroop task (both for the congruent and incongruent tasks).

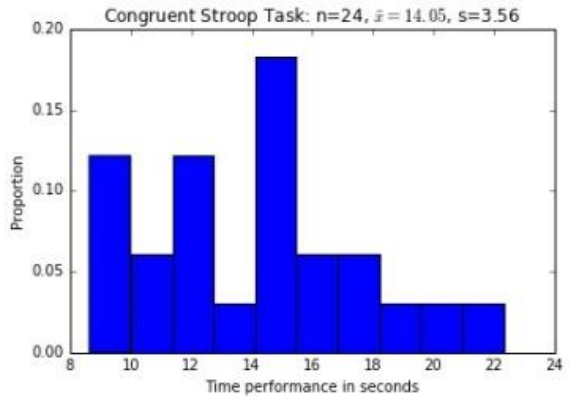
2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.

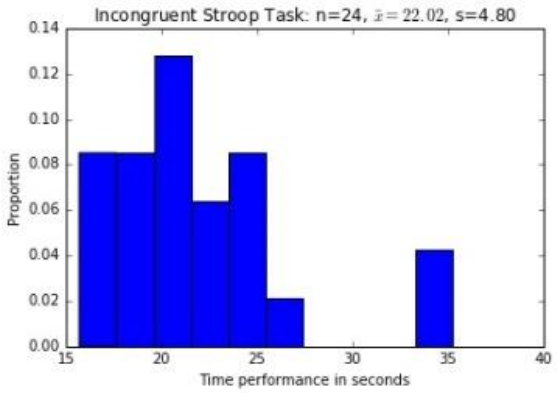
- H_0: $\mu_{congruent} \geq \mu_{incongruent}$; the null hypothesis is that there is no statistically significant difference between the two population means; or else, that the treatment (the incongruent Stroop task) actually improved performance in recognizing colors visually against misleading text. This means that incongruence between color names and colors should not worsen performance in the task.
- H_A: $\mu_{congruent} < \mu_{incongruent}$; alternatively, the treatment worsens people's abilities, on average, to recognize colors visually when confronted with misleading text.
- We choose to perform a one-sided (lower-tailed) t-test, in which we expect that the mean difference will be non-trivial and negative if we are able to reject the null hypothesis. We have chosen a t-test as we do not have access to population parameters μ and σ, making this the best choice to infer something about the population mean difference. Further, we assume the distribution of performance times to be approximately Gaussian normal, due to random differences in performance ability across the population.
- This is a two-condition dependent samples analysis, as the subjects across the two conditions are the same, and as performance on the first task almost certainly has some bearing on performance of the second.

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

- Sample means rounded to 2 decimal places: $\bar{x}_C = 14.05$; $\bar{x}_I = 22.02$;
- The mean of the differences ($x_c - x_i$): $\bar{x}_d = -7.96$;
- Similarly, sample standard deviations: $s_c = 3.56$; $s_i = 4.80$;
- The standard deviation of the sample differences: $s_d = 4.85$;
- Standard error of the difference: SE = 0.99, with n = 24;

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.





- In these normalized histograms, we see that the mode for the congruent task is about 15 seconds, whereas it is at about 21 seconds for the incongruent task. We also see a larger range of values for the latter, all values within 20 seconds of each other, versus the about 11 second range for the congruent task values. We also see that the sample standard deviations reflect this trend in variation. In both samples, the mode appears relatively close to the mean.
- For the incongruent distribution, we see greater positive skew, while the congruent task distribution appears to be more centered.

5. Now, perform the statistical test and report your results. What are 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?

- Computed t-stat: -8.02; α-level: 0.05;
- Lower-tailed critical-value with df = 23: $t_{0.05} = -1.711$;
- CI 95%: (-9.67, -6.26); 2-tailed p-value < 0.0001;[i]
- Conclusion: We can reject the null hypothesis, given that our calculated t-statistic is more than 2.58 standard errors above the critical value with an α-level of 0.05. That is to say, t-stat > $t_{.095}$, and as our p-value is less than .0001, p-value being a measure of the probability of obtaining t-test results at least as extreme as ours (less than .01%).
- Interpretation: This suggests that the incongruent Stroop task significantly impacted subjects' abilities to recognize colors when forced to process two incongruent sets of information: the color as a written word, and the color visualized as a color. As expected, the treatment substantially slowed down subjects' abilities to complete the task.

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!

- At first glance, it seems that the Stroop task effectively requires subjects to engage in two separate cognitive tasks—one, to engage in reading, and two, to recognize a color. But on further thought, the two tasks seem deeply intertwined, as the outcome of the experiment

suggests that we cannot effectively separate the two (our use of language versus our conscious, lingual identification of something that exists externally). When color and name of color are congruent, we can move quite quickly in identifying the colors by name. When they are incongruent, we seem to experience a substantial lag in our ability to process and identify their true names. This may be due to the fact that we first look to language cues to identify a thing, as if our faculty for language has greater priority in our order of recall, even before what one might expect to be the more immediate input of a visual experience (at least when encountered safely on a computer screen—one suspects an awareness of imminent danger would take priority over one's ability to give a name to the threat, as, say, in the case of a tiger nearby).

- Another way to test the Stroop effect in an expanded form would be to give people images with incongruous names: for example, labeling a tiger as a walrus, and asking subjects to name the animal in question. Further, we could test this against other means of receiving lingual inputs—such as spoken words, whether against visual objects (such as colors) or other sounds (i.e. I say "flute" but the sound heard is that of a drum).

[i] Computed using the Graphpad.com online calculator, with the computed t-stat and df=23.