Nancy Olewnik

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

- The independent variable would be changing the word/color combination. It is the variable we are changing to evaluate how they read the data set.

- The dependent variable is the time it takes to read because it represents the output of the experience.

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

**Q2a: Null and alternative hypotheses are clearly stated in words and mathematically. Symbols in the mathematical statement are defined.**

A hypothesis test is a statistical test that is used to determine whether there is enough evidence in a sample of data to infer that a certain condition is true for the entire population. A hypothesis test examines two opposing hypotheses about a population: the null hypothesis and the alternative hypothesis. The population mean is denoted as, $\mu$ . The population variance is denoted as, $\eth^2$ . The sample mean is denoted as $\bar{x}$ and the sample variance as $s^2$.

- $H_0$: $\mu_c = \mu_i$ … The null hypothesis is that the two response times in congruent and incongruent is the same.

- $H_1$: $\mu_c \neq \mu_i$ … Hypothesis 1 is guessing that the average response time for incongruent is not equal to the average response time for congruent.

- The first and second hypothesis will test if there is a difference in reading time between the incongruent and congruent. The hypotheses are about the response time it takes to name, out loud, the color of the inks in which the words are displayed.

- The statistical test I chose is the two-sided paired t-test. We have less than 30 samples, we don't know the population's standard deviation and we assume that the distributions are Gaussian. I chose the two-sided test because it leads more more accurate and reliable results.

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

|  | Congruent | Incongruent |
|---|---|---|
| Mean ($\bar{x}$ ) | 14.0511 | 22.0159 |
| Median | 14.3565 | 21.0175 |
| Standard Deviation ($\eth$) | 3.5594 | 4.7971 |
| Variance | 12.6690 | 23.0118 |

**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**

### Congruent vs Incongruent comparison



### Congruent vs Incongruent comparison
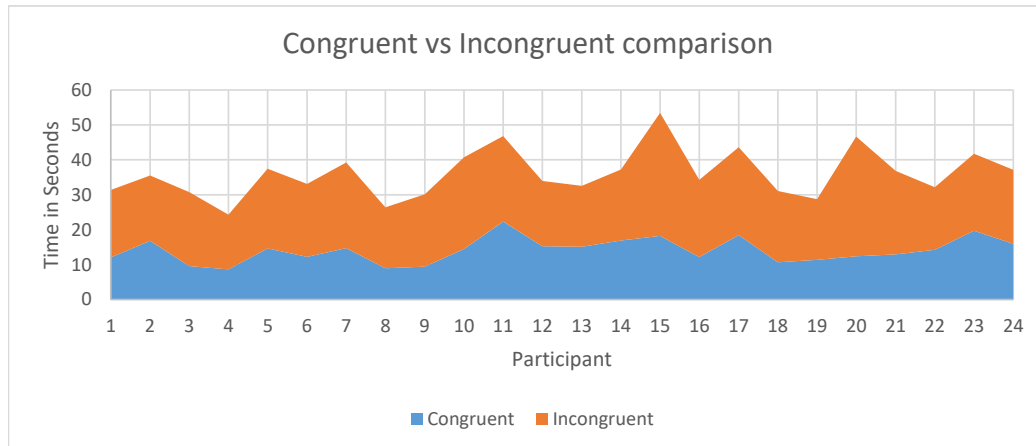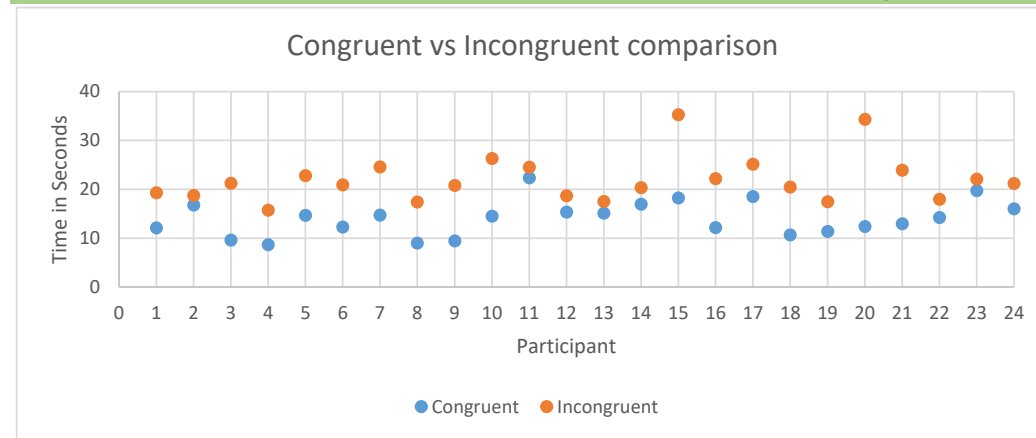


- These graphs illustrate that each participant took longer to read the incongruent data than it took to read the congruent data set.

| | |
|---|---|
| P-value | 0.00000004103 |
| Confidence interval based on 5% confidence level | -10.019 |
| Confidence interval based on 5% confidence level | -5.910555 |
| T-statistic | -8.021 |
| df | 23 |
| T Critical | 2.069 |
| T Critical | -2.069 |

We reject the null hypothesis since the p-value is less than 0.05. This means there is statistical significance of between the response time of incongruent and congruent varies. Based on the confidence intervals, we're 95% confident that the true difference between the congruence and incongruence group average times is The t-critical value is ± 2.069. Our t-statistic is ± 8.021. We just the null because the t-statistic of 8.021 will fall further out. This show statistical significance.
REJECT ... $H_0$: $\mu c = \mu i$ … I will reject the null hypothesis because the average response time varies greatly between the congruent & incongruent data sets.
ACCEPT … $H_1$: $\mu c \neq \mu i$ … The average response time for incongruent data is not equal to the average response time for congruent data

I believe that the effect between the congruent and incongruent data set is dependent on the intelligence level to follow the directions and read quickly. A similar task that would result in a similar effect would be to ask someone to recite the A, B, C's and then ask them to recite the alphabet in backwards sequence, recording the time for each.