

Jaryd Meek
CSCI 3002
Recitation 12
18 April 2020

I compiled the data at approximately 3pm on Friday, April 16. There was one sample that I removed for Prototype A as it had a value of 10000 seconds for task 2 (well above the maximum value allowed for the task), since this is a significant outlier. This lead me to have 83 different points of data to work with, which is a large enough dataset to perform statistical tests on.

I completed a 2 sample T-Test for the difference in means for each of the 3 tasks for the 2 prototypes. After completing the tests, I found the p-value for Task 1 was 0.54, the p-value for Task 2 is 0.48 and the p-value for task 3 was 0.80. Operating under the basis that the null hypothesis is that there is a significant difference in the time taken to complete the tasks on Prototype A vs B, using an alpha level of 0.05 (95% confidence) there is insufficient evidence to reject the null hypothesis. Therefore, there is no statistically significant difference in time that it took participants to complete the tasks on Prototype A vs Prototype B.

Participants responded significantly more positively to Prototype A. Participants found it almost twice as effective as well as significantly more useful than Prototype B. However Prototype B was found to be slightly more fun than Prototype A. People also marked Prototype B as confusing significantly more often than Prototype A. Overall, the population significantly preferred working with Prototype A instead of Prototype B.

Overall I would **greatly** recommend Prototype A over Prototype B. Prototype A may have been a bit more boring overall than Prototype B, but the text was significantly clearer, as well as the options being laid out in a easier to navigate order than the circles. Also from an accessibility perspective, Prototype A is significantly more accessible, and significantly easier to adapt to allow anyone to use.