HCI Assignment 6: User Study Evaluation

Daniel Leal and Fabian Gubler

December 2022

1 Group Design

1.1 Experiment Setup

1. For our study, we invited 20 participants from your target group into your lab.
2. Every participant completes the same experiment experiment by using the old version of your application.
3. In addition participants fill out a System Usability Scale (SUS).
4. Once they are finished, all participants continue with the new version.

1.2 Paired and Unpaired Data

The collected data comes from one sample where every participant has done all conditions (i.e. Paired Samples).

2 Variables

2.1 Independent Variable

Our independent variable is the design of the checkout process. In its old form, the customer had to input every detail about their delivery address manually when ordering their food. With the redesign using the address lookup API we are hoping to optimize our desired effect by changing the cause (i.e. independent variable)

- The independent variable is the cause.
- Its value is independent of other variables in your study.

2.2 Dependent Variable

We are measuring the completion time of each participant during ordering process. Minimizing the time of completion is the desired effect of our new changes to the food delivery application.

- The dependent variable is the effect.
- Its value depends on changes in the independent variable.

3 Completion Time Measurment

3.1 Data Category of Completion Time

Completion time data falls into the category of quantitative data. In particular, it corresponds to ratio data, as has a true and meaningful zero (i.e. no negative time) and is not limited to a specific interval.

3.2 Was the re-design of the application successful?

```
39 import pandas as pd
38
37 df_new = pd.read_csv("/home/fabian/nextcloud/mcs/
36 df_old = pd.read_csv("/home/fabian/nextcloud/mcs/
35
34 print(df_new.describe())
33
32 # count         20.000000
31 # mean         150.550000
30 # std           21.663637
29 # min          120.000000
28 # 25%          133.500000
27 # 50%          149.000000
26 # 75%          164.250000
25 # max          190.000000
24
23 print(df_old.describe())
22
21 # count         20.000000
20 # mean         169.550000
19 # std           24.686402
18 # min          133.000000
17 # 25%          152.500000
16 # 50%          162.500000
15 # 75%          192.000000
14 # max          206.000000
```

Figure 1: Basic Statistical Measures of both old and new application

Our simple analysis shows the on average the new application checkout process could be done in smaller amount of time (150.55 vs. 169.55). In particular, users could complete on average about 20 seconds faster. The newer application is

also more consistent in terms of how the variance of the data is (smaller standard deviation for newer application). Both datasets do not have notable outliers, when looking at maximum and minimum values.

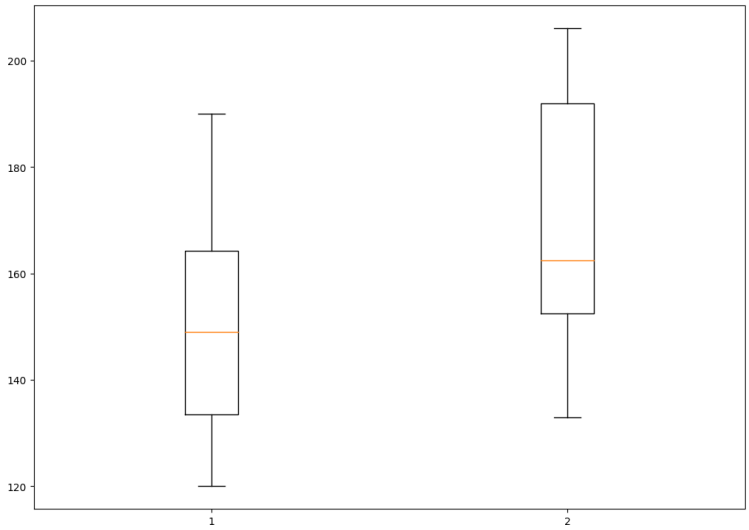

Figure 2: Box Plot of both data sets side by side for comparison

By looking at the boxplot of both datasets (left = new application, right = old application) we can again see the improvements of the redesign. We decided to use a boxplot for plotting the data, as we can see various aspects which we can immediately compare of both sets. First, we see that the spread of data is similar, but the respective minimum and maximum entries differ (i.e. redesign has better result). In addition we can see that the median of the redesign is fairly in the middle of the samples (approx. 150). In contrast in the case of the old, we see that the median value is lower in comparison (approx. 160). As a consequence, if we take the median (instead of the mean in the above analysis) as our final measure of success, we get a result that is not that significant. Hence, when also taking into account the small sample size, we cannot conclude a significant improvement.

4 Improvements to the Design

4.1 What could be improved about the experiment design and why?

The initial goal of the redesign, was to make the checkout process "easier and faster". The customer burden of having to input every detail on their own, would suggest that she/he might be annoyed or even frustrated. In addition to this quantitative approach, we would suggest to add a more quantitative approach, where the participant would be asked several questions. Introducing quantitative elements to the experiment is particularly feasible, as the participant where in-person already, resulting in the chance to observe them and pose questions directly during and after use.

4.2 What is another measure you could take for your experiment?

When restricting to one measure an improvement would be to measure the **amount of button clicks**. In contrast to completion time, this still will measure the perceived speed but also focusses on the simplicity and ease of use.

Focus should be whether the process was seamless or frustrating. Also one could ask the participant how much value she/he lays on the speed and ease of the checkout process. Also based on the fact that the experiment showed not a significant improvement, one could ask the participant which design he prefers (manual input, vs. automatic lookup API).