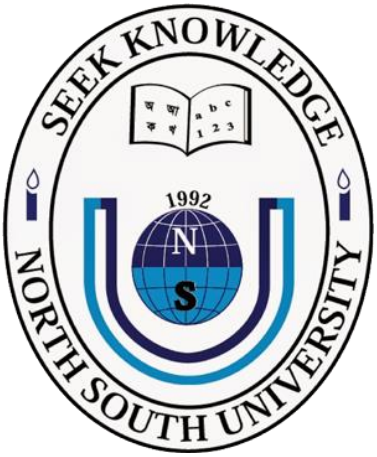


Report on

"Analysis of Text Entry Performance Metrics Using SwiftKey Keyboard"

Submitted To: Ms. Tamanna Motahar

Course: CSE299

Date: 17/09/19

Name	ID
B. M. Raihanul Haque	1512756042
Md. Nazmul Islam	1511410042
Mahmudunnabi	1511264042

Table of Contents

Abstract: Many researchers, throughout the years, suggested various text entry methods to assist the users to type faster and precisely. However, results obtained from these methods are vague and significantly diverse from one another and it is difficult to extract proper meanings from the values of entry speed and error rate. WebTEM, developed by Dr. Ahmed Sabbir Arif who is an assistant professor of Electrical Engineering and Computer Science at the University of California, is a web application to record text entry metrics. This application allows the researchers to get proper values and make sense out of it. SwiftKey keyboard, an application available for android device, provides multiple keyboard layouts. WebTEM application is proper platform to analyze error and entry rate of users who have used these layouts.

Experimental Design: Our objective of the experiment is to determine if we are able to detect measurable differences in dependent variables between levels of the independent variables.

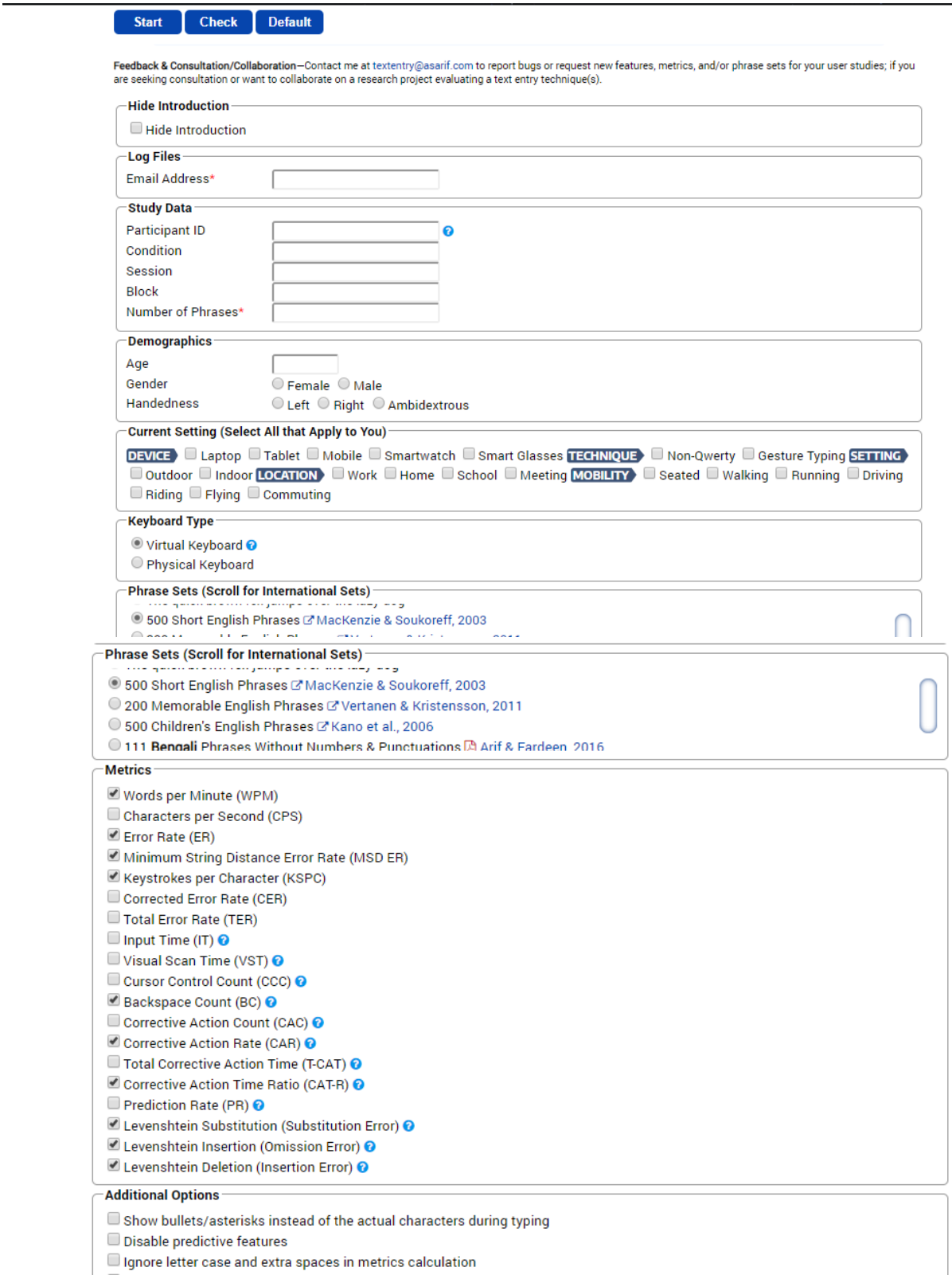
1. **Setup:** The experiment was set up in a way so that it falls under within-subject design where each participant was exposed to multiple conditions. The keyboard layouts are:
    i. Default keyboard layout
    ii. High contrast keyboard layout
    iii. Light keyboard layout

    Multiple conditions along with the layout includes:

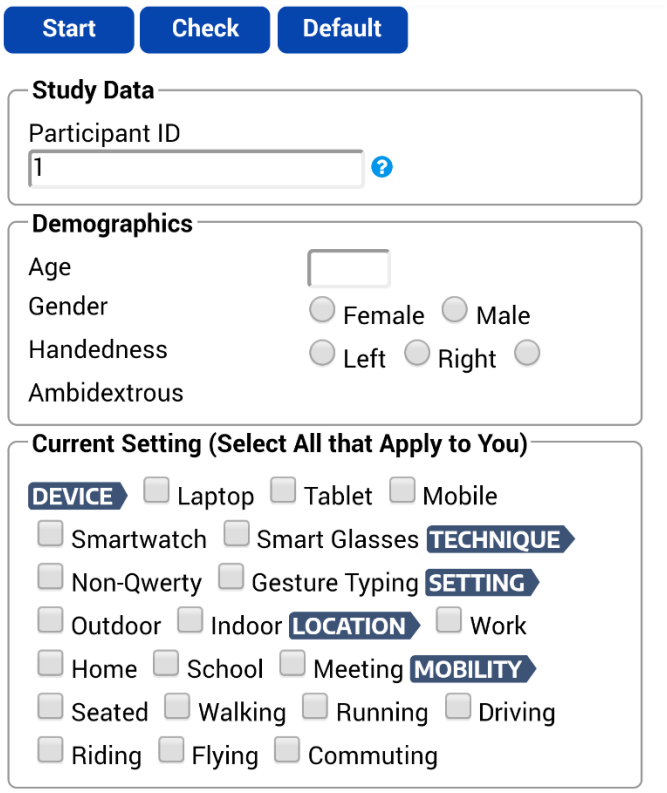
    i. Mobility (Seated/Walking)
    ii. Device (Laptop/Tablet/Mobile)
    iii. Setting (Outdoor/indoor)
    iv. Location (Home/School)
    v. Handedness (Left/Right/Ambidextrous)

    Here, our independent variables (inputs, they are what we manipulate) are the keyboard layouts and all other conditions imposed on the participants. Dependent variables (outputs, they are what we measure) are total error rate and entry rate (word per minute).

2. **Participants:** All the participants were students; either university or college student. Both male and female participated as part of the experiment. They were randomly selected and no particular groups were formed. All of them were ambidextrous and their ages fall in between 18 to 25 years.
3. **Apparatus:** The components that were used to conduct this experiment are:
    i. WebTEM Application
    ii. SwiftKey Keyboard
    iii. Mobile device



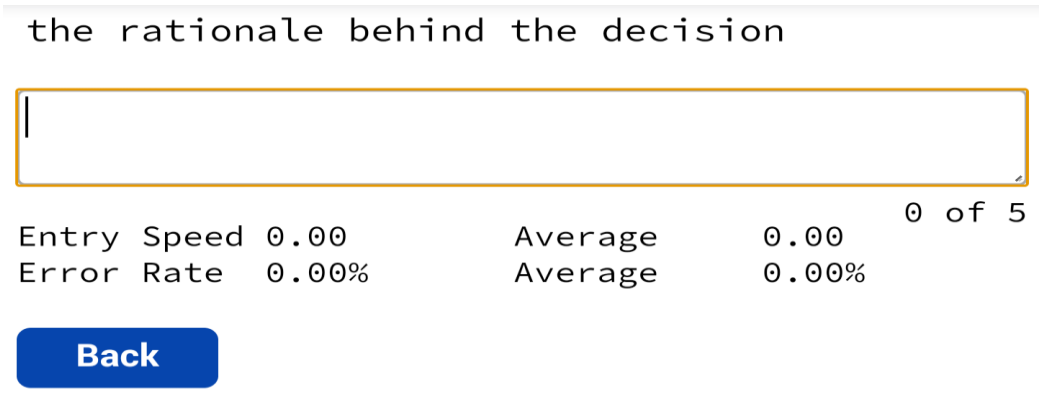
Interface for setting up the link of different layouts



Interface of WebTEM Application (after the link has been set up)

Three different layouts (Default, High Contrast, Light)

4. **Material:** Each session contained a phrase set of randomly generated 5 sentences provided in English language.



A randomly generated sentence from phrase set

5. **Procedure:** Each participant was enlightened about the survey that they were going to take and getting the verbal agreement from them the survey starts by filling out the necessary information. All the survey is done within the indoor compound of either a university or a home. Each participant had to type 30 sentences. There were certain order keyboard layout and mobility which are:
    i. Default seated (5 sentences)
    ii. High contrast seated (5 sentences)
    iii. Light seated (5 sentences)
    iv. Default walking (5 sentences)
    v. High contrast walking (5 sentences)
    vi. Light walking (5 sentences)

    When the survey is done, this application generates one information (.INFO) and two tab-delimited files (.TSV) (this one can be represented as an Excel sheet) to record all settings, timestamped events, and performance metrics, respectively. Phrases are represented by each row of the metrics log and a metric is represented by each column. The last row contains the average values.

    There are three error correction conditions:

    a. None: Participants are not allowed to correct errors.

    b. Recommended: Participants are recommended to correct errors as they identify them.

    c. Forced: Participants are forced to correct each error.

    We imposed the 2nd condition on the user that's why Total ER is used to measure error rates. For WPM, no such condition is required.

Data analysis: There were 10 participants and 50 data values for each unique setting of layout and mobility. Using the values found in WPM and Total Error Rate column we have been able to find the mean and standard deviation and plot histogram.

Total Error Rate

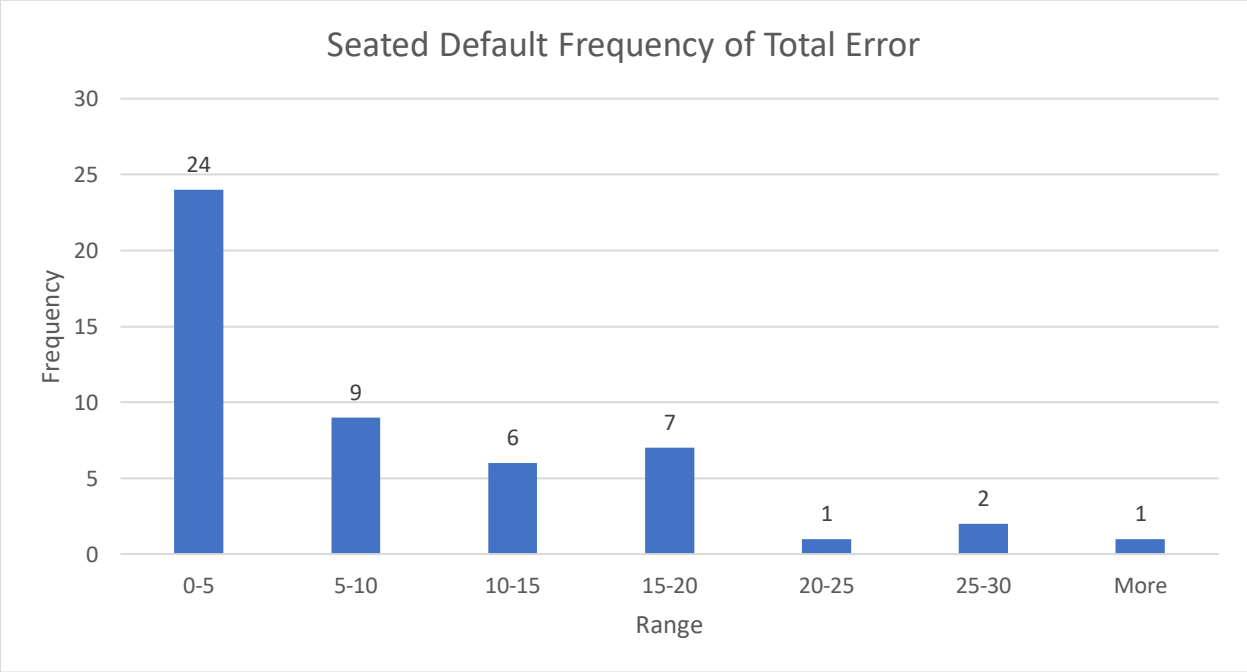
The formula that is used to calculate the total error is:

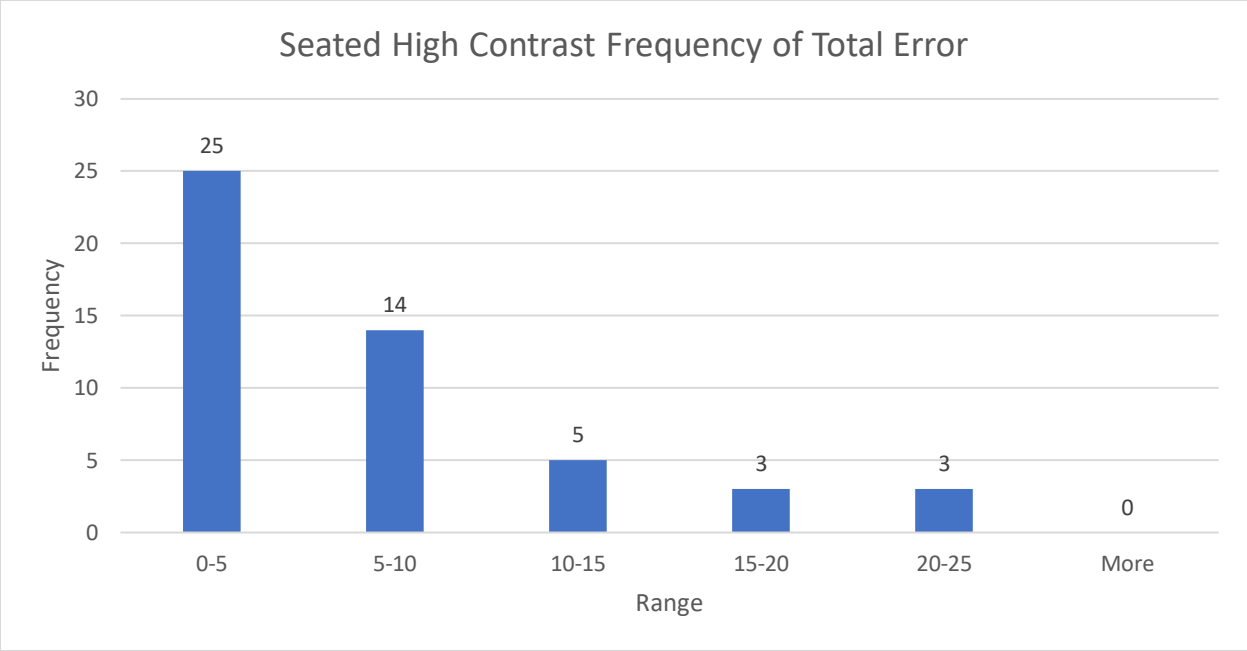
$$Total\ ER = \frac{INF+IF}{C+INF+IF} \times 100\%.$$

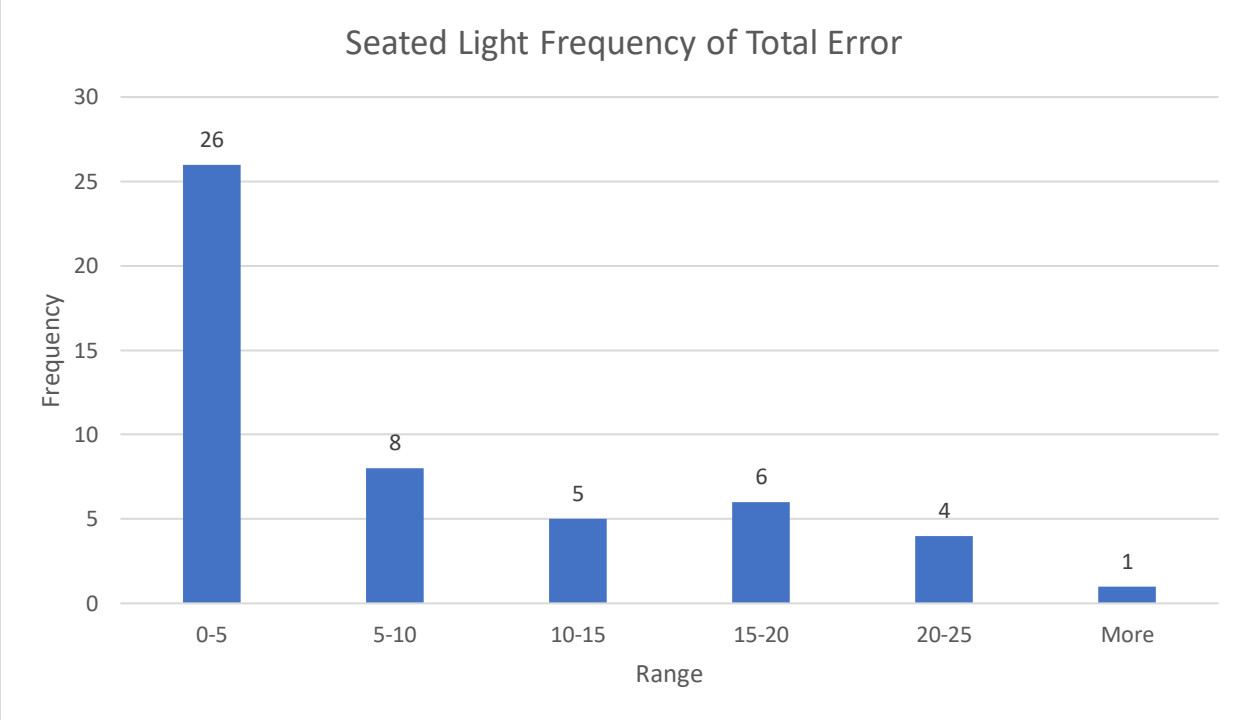
Where, INF = is the number of unnoticed errors (incorrect characters),

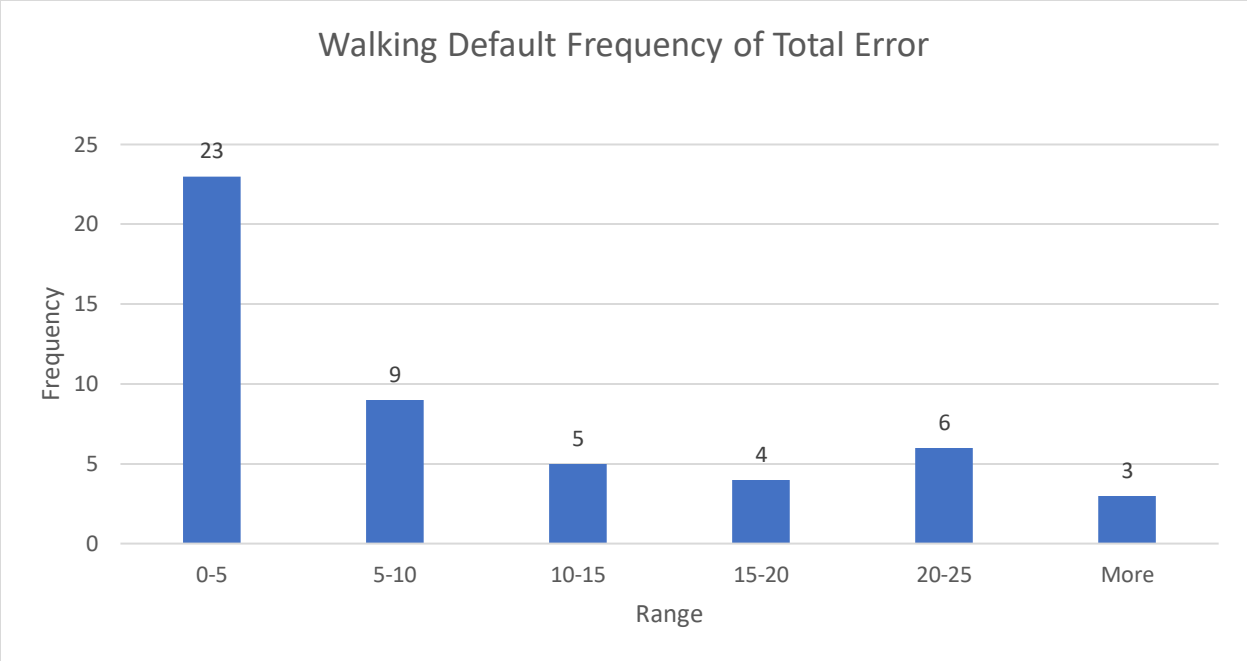
IF = keystrokes that are not editing keys (backspace, delete, cursor movement), which do not appear in the final transcribed text result,

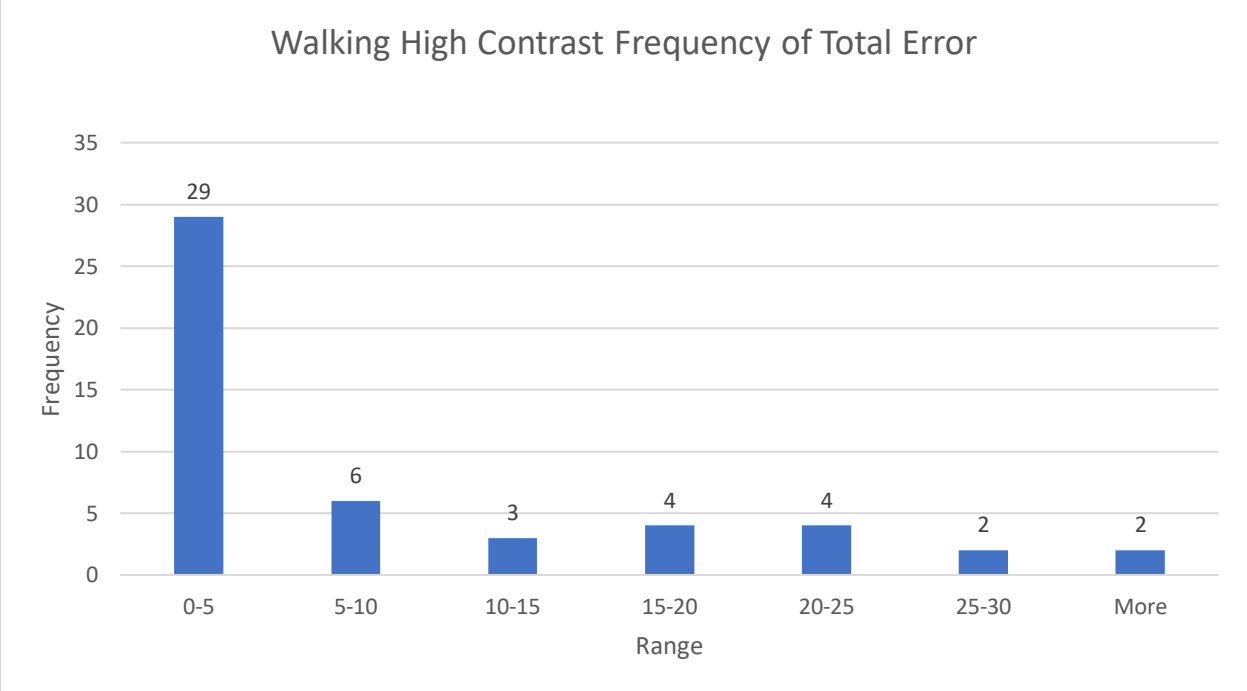
C = is the number of correct characters in the transcribed text.

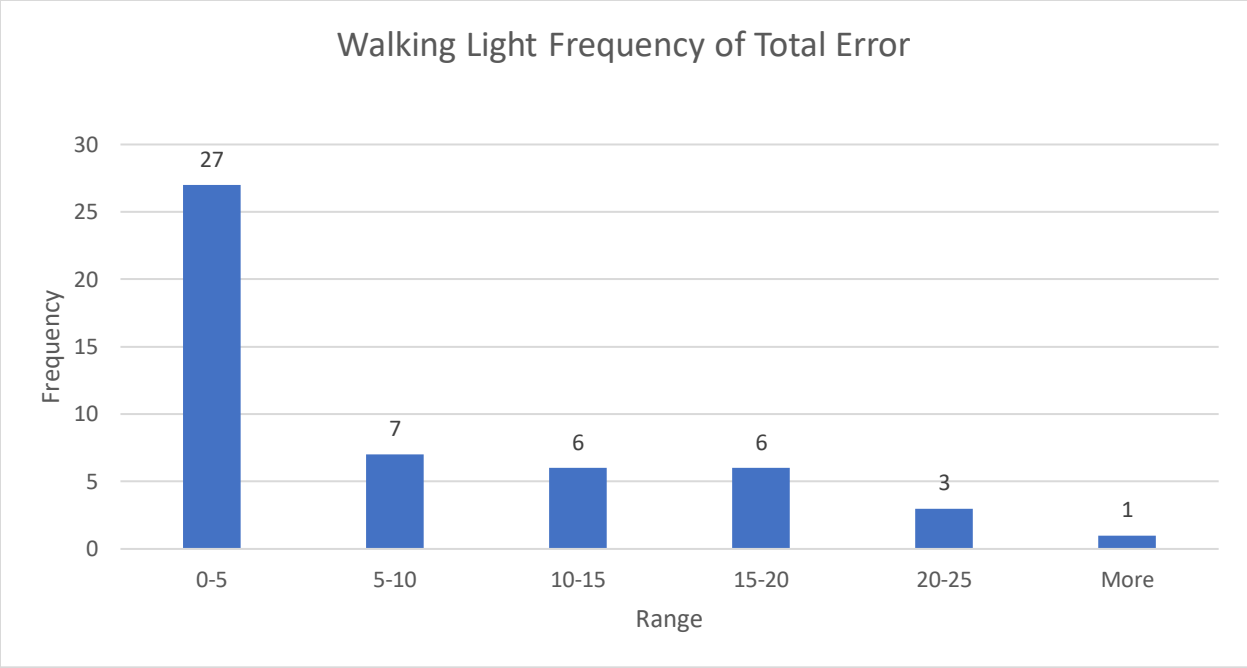



Seated High Contrast Frequency of Total Error
30
25
20
15
10
5
0
Frequency
25
14
5
3
3
0
0-5
5-10
10-15
15-20
20-25
More
Range


Seated Light Frequency of Total Error
30
25
20
15
10
5
0
Frequency
26
8
5
6
4
1
0-5
5-10
10-15
15-20
20-25
More
Range


Walking Default Frequency of Total Error
25
20
15
10
5
0
Frequency
23
9
5
4
6
3
0-5
5-10
10-15
15-20
20-25
More
Range


Walking High Contrast Frequency of Total Error
35
30
25
20
15
10
5
0
Frequency
29
6
3
4
4
2
2
0-5
5-10
10-15
15-20
20-25
25-30
More
Range



All these histograms suggest that, more than 50% of the time there exists less or no error in user typed sentences.

Word Per Minute (WPM)

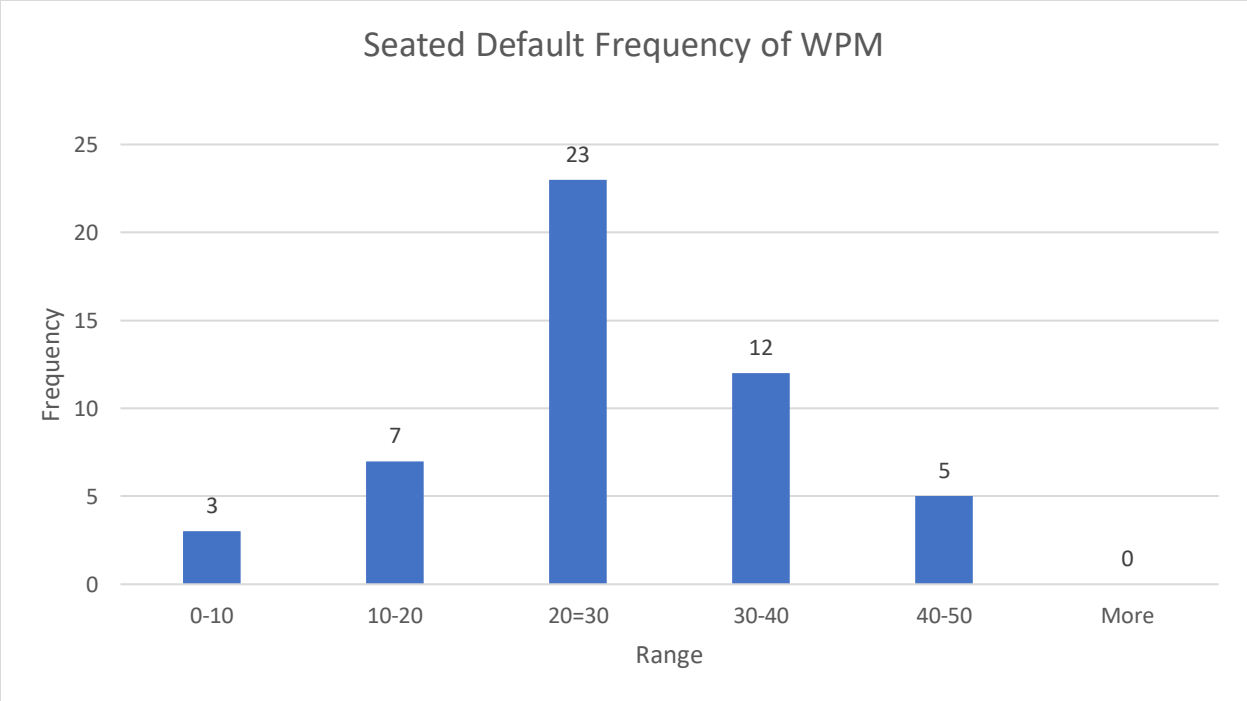
The formula that is used to calculate the total error is:

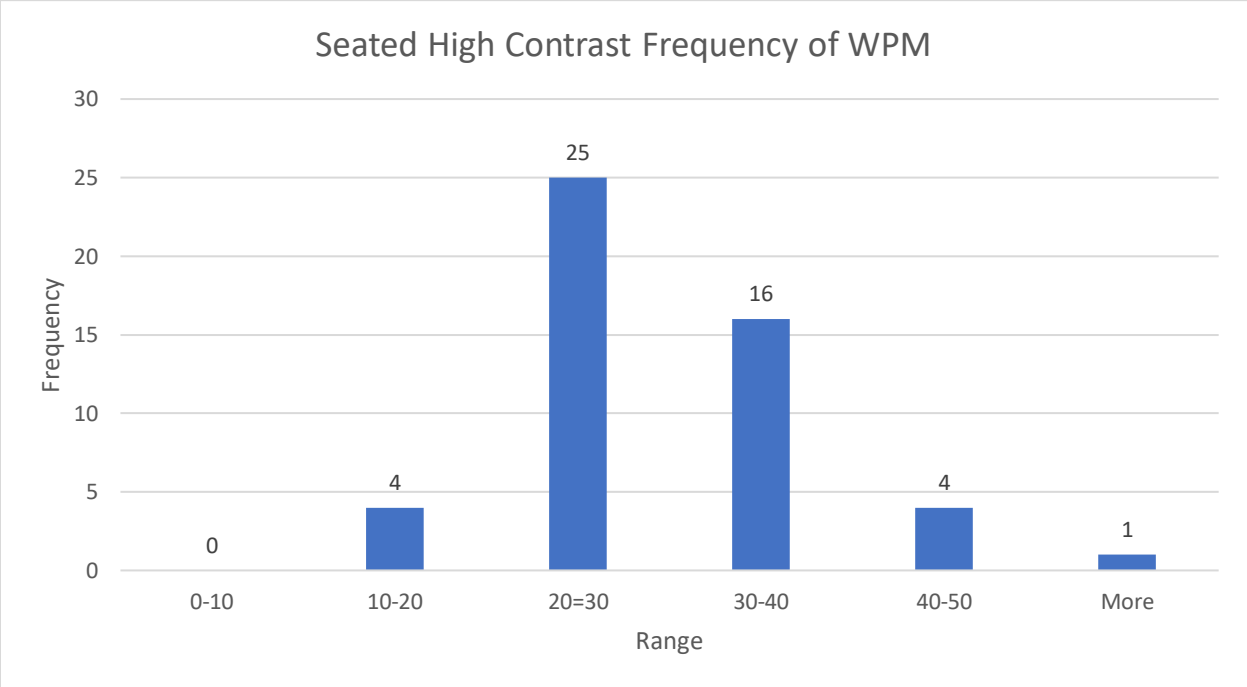
$$WPM = \frac{|T|-1}{S} \times 60 \times \frac{1}{5}.$$

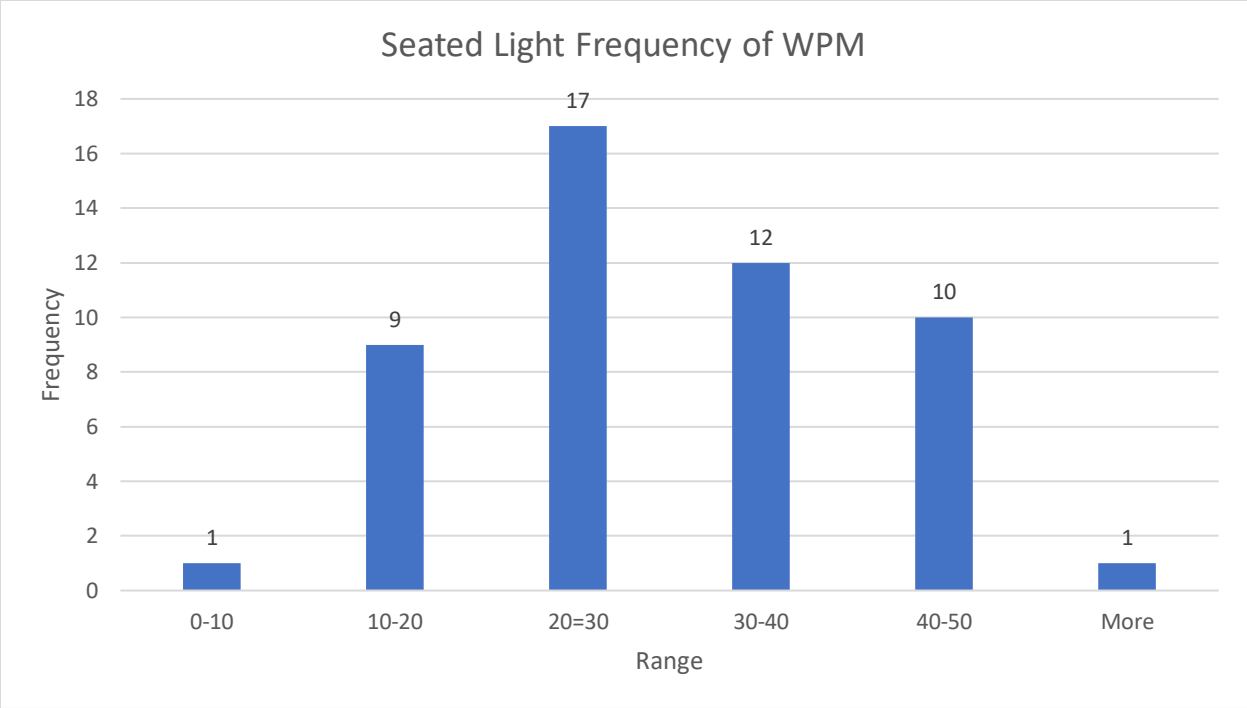
Where, |T| = length of the transcribed text,

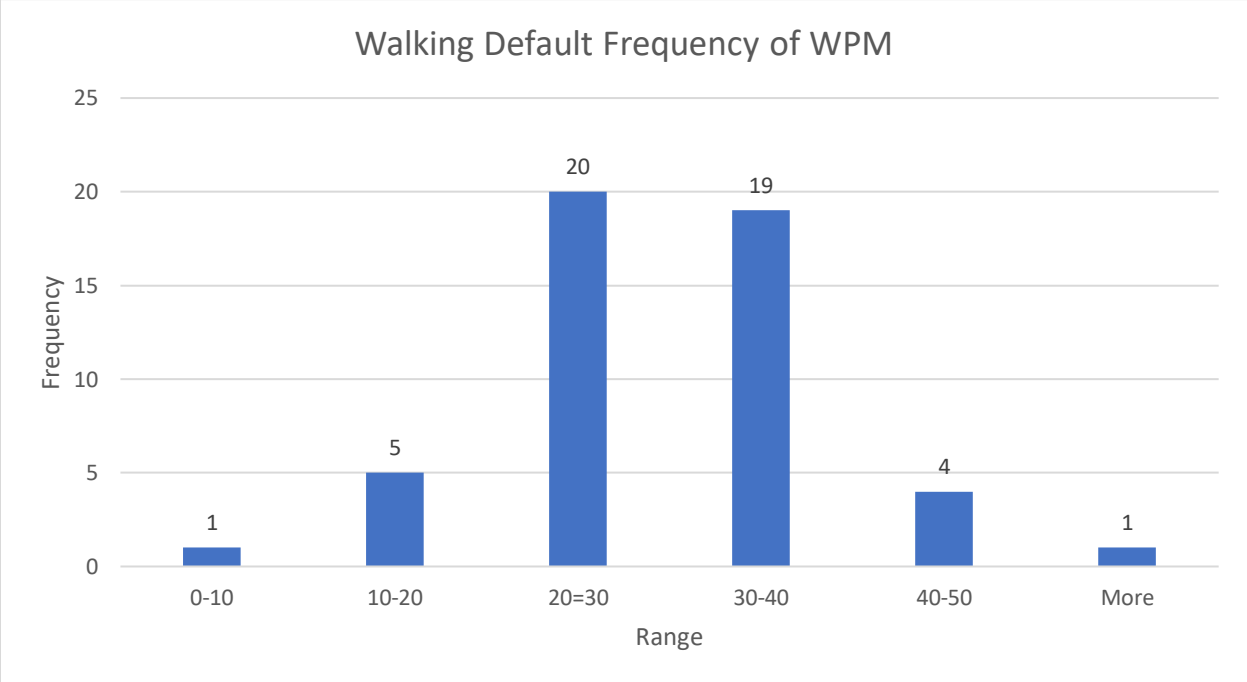
S = time in seconds measured from the first key press to the last,

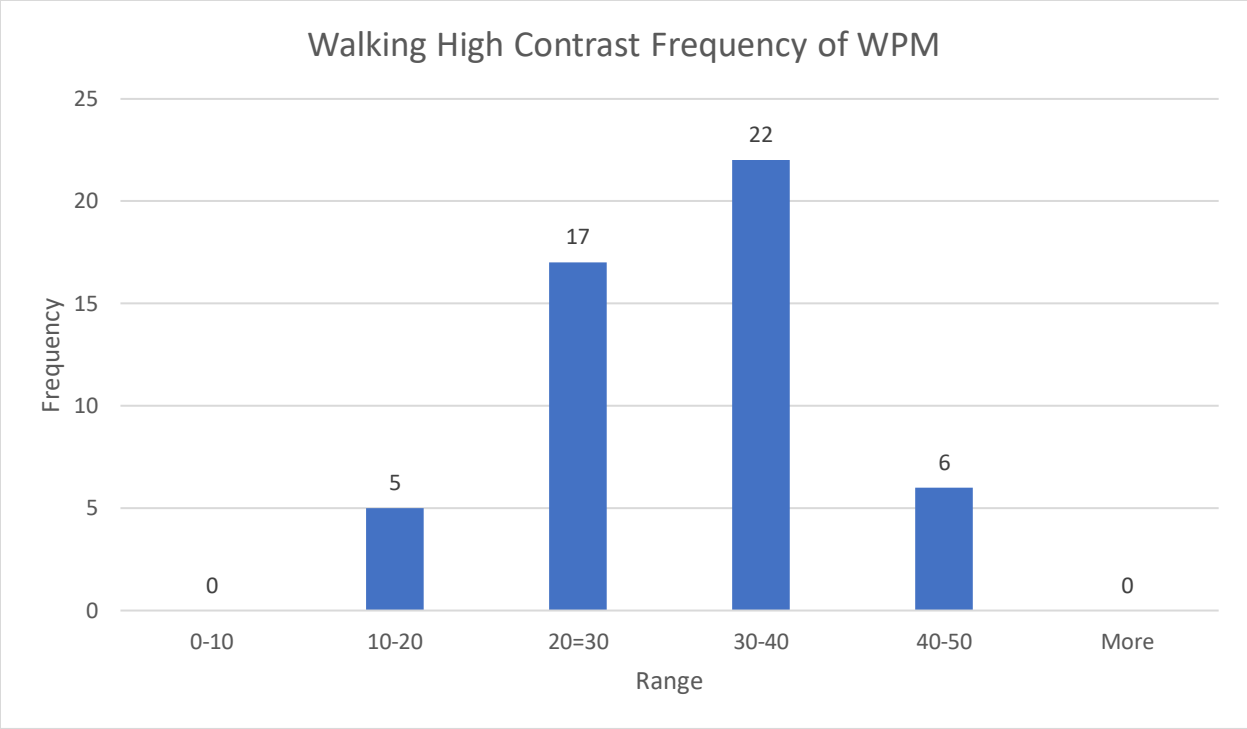
The constant 60 is the number of seconds per minute, and the factor of one fifth represents the average length of a word in characters.

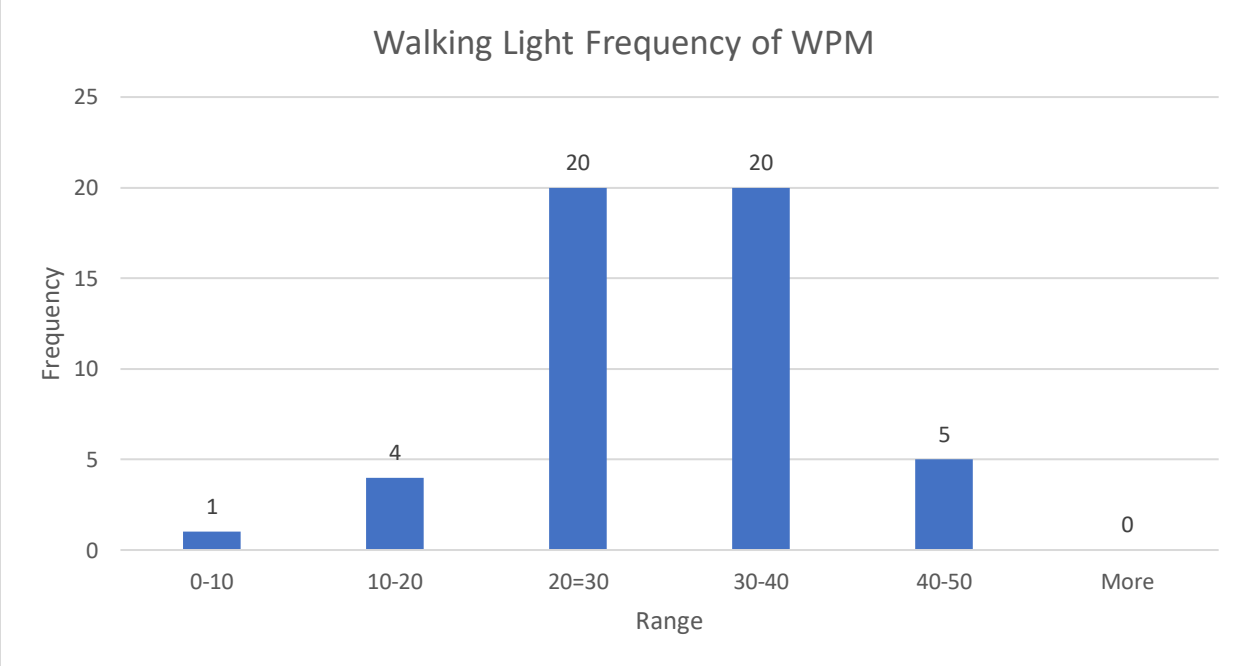

Seated Default Frequency of WPM
25
20
15
10
5
0
Frequency
23
12
7
5
3
0
0-10
10-20
20=30
30-40
40-50
More
Range


Seated High Contrast Frequency of WPM
30
25
20
15
10
5
0
Frequency
25
16
4
4
1
0
0-10
10-20
20=30
30-40
40-50
More
Range


Seated Light Frequency of WPM
Frequency
18
16
14
12
10
8
6
4
2
0
1
9
17
12
10
1
0-10
10-20
20=30
30-40
40-50
More
Range


Walking Default Frequency of WPM
Frequency
25
20
15
10
5
0
1
5
20
19
4
1
0-10
10-20
20=30
30-40
40-50
More
Range





All these histograms formed bell shaped curve which indicate that, the typing speed is neither too fast nor too slow rather the mean is a good approximation for the center of the data.

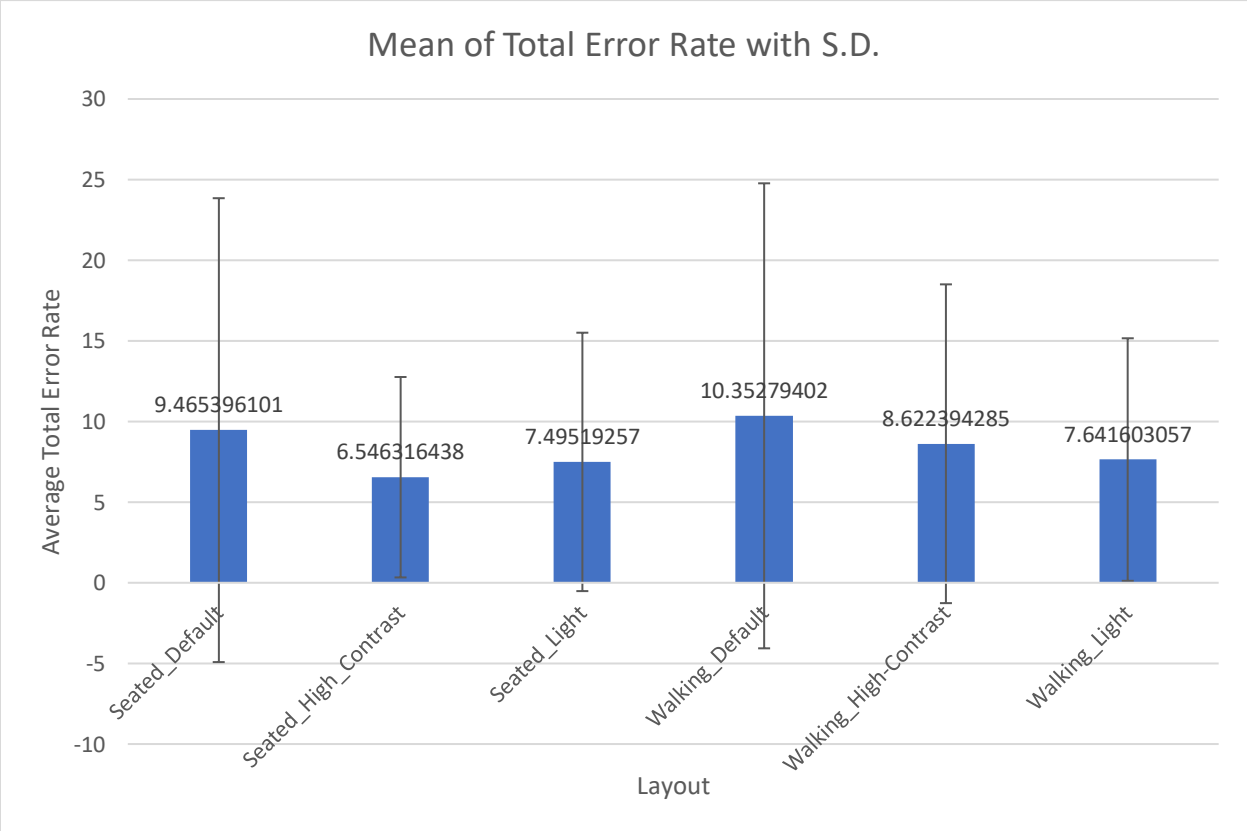
The Null Hypothesis and Result: Before we began our experiment, we have initiated a null hypothesis and an alternate hypothesis for total error rate and we did the same for WPM.

Hypothesis for Total Error Rate

• **H0: Accept the null hypothesis.** There exists less error for a person typing while seated using any keyboard layout rather than people walking and typing using the same layout.

• **H1: Reject the null hypothesis.** There exists equal or more error for a person typing while seated using any keyboard layout rather than people walking and typing using the same layout.

Layout	Mean	Standard Deviation
Seated_Default	9.465396101	14.38282558
Seated_High_Contrast	6.546316438	6.21446522
Seated_Light	7.49519257	8.011682326
Walking_Default	10.35279402	14.418035
Walking_High-Contrast	8.622394285	9.883688641
Walking_Light	7.641603057	7.520519885



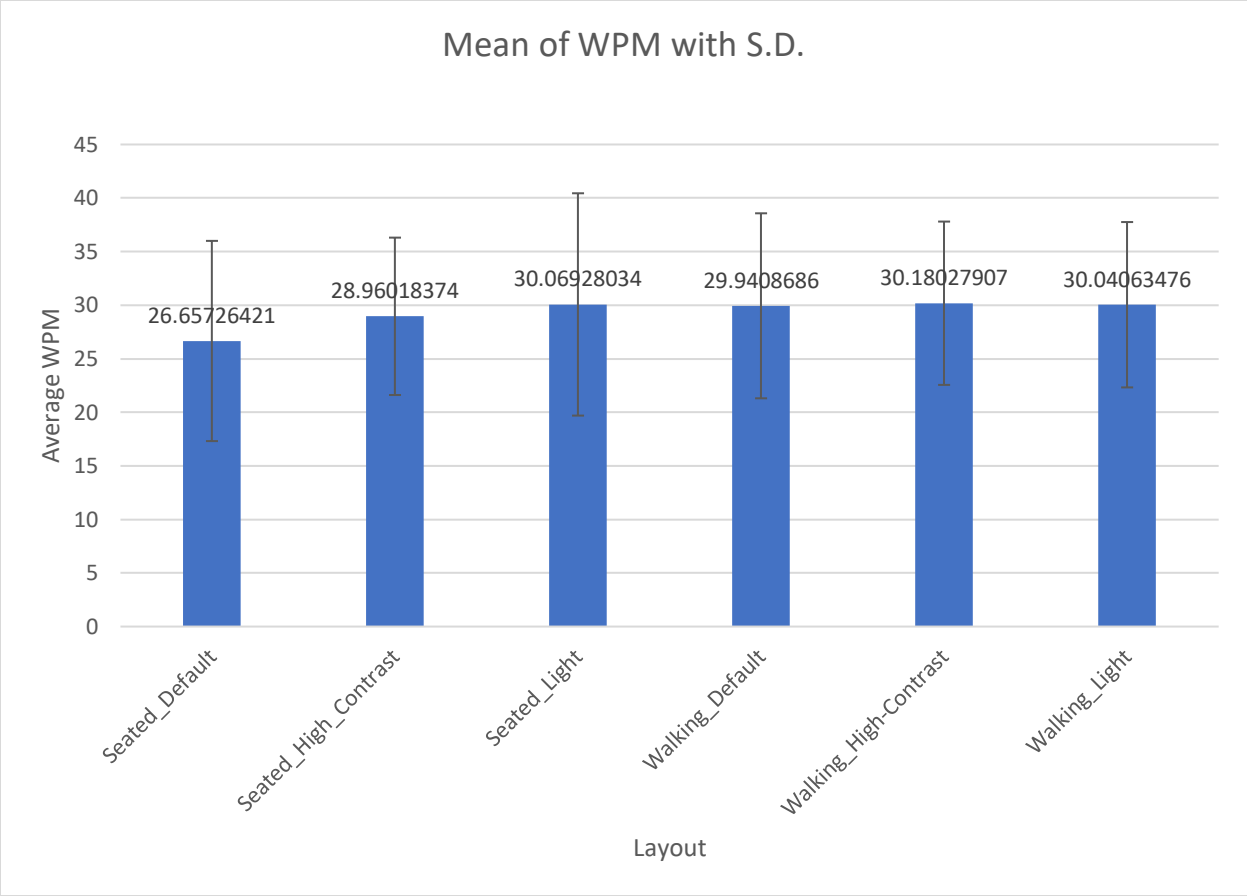
The data helps us to decide to accepts the null hypothesis and rejects the alternate hypothesis. Thus, less error exists for a person who is typing while seated using any keyboard layout rather than people walking and typing using the same layout.

Hypothesis for WPM

• **H0: Accept the null hypothesis.** People typed faster while seated using any keyboard layout rather than people walking and typing using the same layout.

• **H1: Reject the null hypothesis.** People's typing speed is equal or less while seated using any keyboard layout rather than people walking and typing using the same layout.

Layout	Mean	Standard Deviation
Seated_Default	26.65726421	9.344741724
Seated_High_Contrast	28.96018374	7.346082008
Seated_Light	30.06928034	10.37621813
Walking_Default	29.9408686	8.634826718
Walking_High-Contrast	30.18027907	7.622438403
Walking_Light	30.04063476	7.718105947



The data allows us to nullify the null hypothesis and accepts the alternate hypothesis. Thus, people typing speed is equal or less while seated using any keyboard layout rather than people walking and typing using the same layout.

This also proves that faster typing causes more errors.

The overall experiment tells us about one of the key features of SwiftKey keyboard. The light keyboard layout is a reliable layout as it produced less error and was comparably faster.

Limitations and Future Scope: Although this was a controlled HCI experiment yet there was limitation to our experimental design. Since, this was a within-subject design we were required to avoid skill transfer effects which is, participants taking part of an experiment tend to get better at performing the task as the experiment progresses. As the order of layout arrangement and mobility was same this might affect our data.

We have only analyzed total error rate and WPM. But there is more to it. The same web application can be used for, KSPC Analysis, EKS ER Analysis, ER and the MSD ER Analysis, Visual Scan Time Analysis and so on.

Conclusion: From the experiment we have got many significant results, like people are more efficient at typing when seated and SwiftKey light keyboard layout is comparably better than other layouts that have been tested. This experiment helped us to understand the typing behavior of the individual and how they can get better and which of the SwiftKey layout can make typing more efficient.

References:

1. Statistical Analysis by Per Ola Kristensson, Department of Engineering, University of Cambridge
2. Analysis of Text Entry Performance Metrics by Ahmed Sabbir Arif and Wolfgang Stuerzlinger, Dept. of Computer Science & Engineering, York University, Toronto, Canada
3. WebTEM: A Web Application to Record Text Entry Metrics by Ahmed Sabbir Arif and Ali Mazalek, Ryerson University, Toronto, Ontario, Canada