Comparing the Typing Performance of Traditional Typing and Swipe Typing

Nguyen Minh Sang 118031 Erich Flock 117954 Fan Fan 118043
sang.nguyen.minh@uni-weimar.de erich.flock@uni-weimar.de fan.fan@uni-weimar.de

ABSTRACT

Mobile devices are widely used to type texts in a large variety of contexts, like sending an email or writing a comment in Facebook. Knowing the importance of typing on mobile devices, we decided to compare the performance between two different typing methods: traditional typing and swipe typing. We conducted an experiment with 12 participants who were asked to perform different tasks involving typing while we tracked the time. Later, we analyzed all the data to verify our hypothesis: Swipe typing is faster than traditional typing.

INTRODUCTION

Swipe typing technology is provided by a company named SwiftKey. This technology allows users to type one word without removing the finger from the keyboard by swiping from one letter to another. This feature suggests that a user would type faster by swiping through the letters than by typing each letter of a word individually. Our experiment was designed to test this statement in the context of mobile phones. Firstly, we designed the test plan. Then, we created an app which would be used during the user study for the users perform the tasks and also to track the time in each task. After collecting all the data from the participants, we analyzed it and we used the results to verify if swipe typing is really faster than traditional typing.

RELATED WORK

We started our research by searching published papers that tried to compare the performance of these two different methods of text input. We first analyzed the paper of a user study applied by researcher from Misericordia University of Dallas, USA [1]. In this paper, the authors compare the performance of swype and conventional keyboards in a desktop context. The researchers concluded that there is no significant difference of performance between these two keyboards, although the participants preferred the typing by swiping. The second article we analyzed was written by the Department of Computer Science and Engineering of York University, Toronto, Canada [2]. In this article, the authors compared four variations of QWERTY keyboards for mobile phones: standard QWERTY keyboard, octopus keyboard, TouchPal Curve keyboard (it has the same technology as SwitfKey, which allows the user to swipe through the letters and form one word), and TouchPal T + keyboard. In this study, the researchers asked the users to type the same phrase 8 times for each type of keyboard and then, they asked the users about the best and the worst one. Surprisingly, the Curve keyboard was top voted as the worst one and had the worst performance among the keyboards. These studies motivated us to conduct our one user study to verify the performance of two types of typing.

RESEARCH QUESTION AND HYPOTHESIS

In our empirical study, we want to evaluate the performance of typing using the traditional method and the swipe method. The main research question of our study is:

- Is swipe typing faster than traditional typing on mobile phones?

Based on this question, we formulated one hypothesis:

- H0: Swipe typing is faster than traditional typing when the user is sitting.
- H1: Swipe typing is faster than traditional typing when the user is walking

DATA COLLECTION

With the hypothesis defined, we have the following independent and the dependent variables:

- Independent: Swipe typing, Traditional typing, Sitting, Walking, Device

- Dependent: Task completion time

We designed the experiment in order to reduce the influence of external variables, so we used the same mobile phone for all the sessions and also alternated the order of the tasks in each session. To check the speed of typing, we collected the time of each task completion in milliseconds. The same text is used for all the tasks.

The user study sessions were not conducted in a laboratory but on the field. We preferred to apply the experiment in a context where the users would normally type during their daily activities.

We built a custom app to use in our experiment. This gave us freedom to design a good test plan and also provided us the chance to get quantitative data directly from the application (time tracking).

Before each session, we showed the user how swipe typing works and we let them try it before beginning the session. This reduces the influence of experience factor, which we could not completely control and is a limitation of our experiment.

EXPERIMENT PREPARATION

In order to conduct our user study and collect the time of each task, we created an Android application with four different tasks (Fig. 1):

- Typing a text using the traditional text input while sitting
- Typing a text using the traditional text input while walking
- Typing a text using the swipe text input while sitting
- Typing a text using the swipe text input while walking

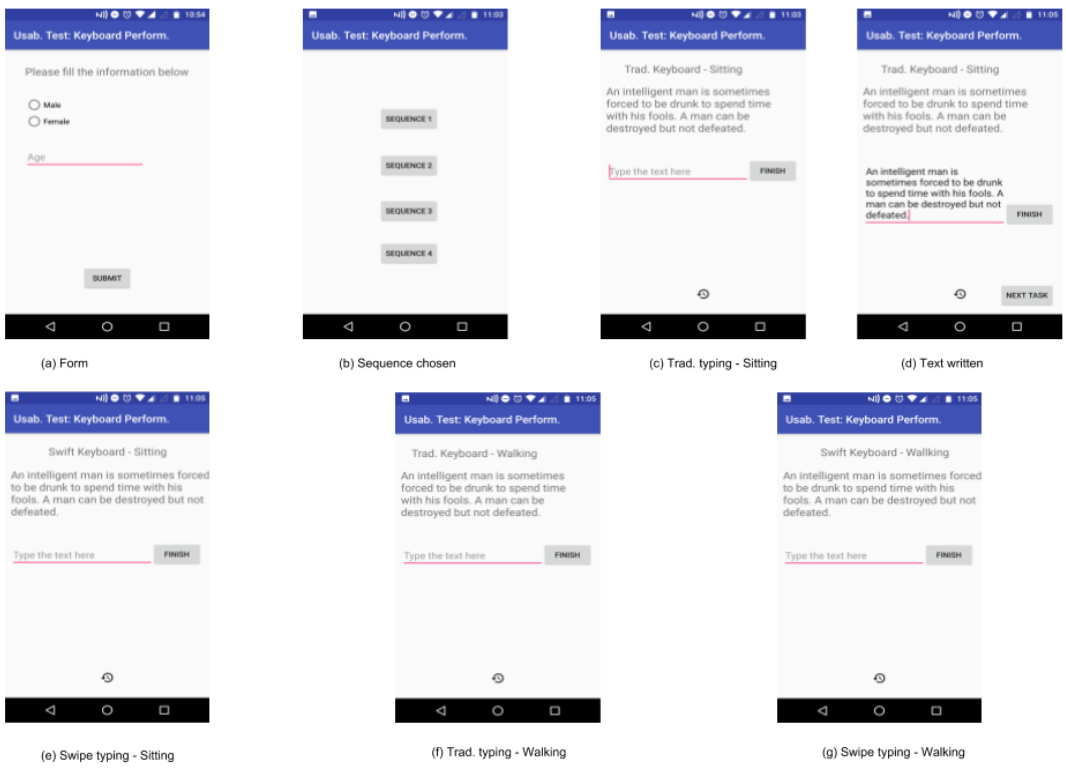

Fig. 1. Overview of the Android App screens and their functions.

These tasks were alternated in 4 different sequences to avoid the influence of learning during the sessions.

In each task the users had to write the same text. The text was carefully selected to contain a good number of words (not too many, not too few) and also a widely variety of characters, but it also had to exclude the two characters “y” and “z” because to avoid the problem of difference between

German and non-German layout: *"An intelligent man is sometimes forced to be drunk to spend time with his fools. A man can be destroyed but not defeated."* This text is a combination of two different quotes by Ernest Miller Hemingway (July 21, 1899 – July 2, 1961), an American novelist.

As we did not find an application that we could use to conduct our user study, we decided to create our own application. We chose to create an Android app because it is the most used operating system among mobile phones (87.7% of world market share in 2017 [3].

The first screen (Figure 1, element (a)) has two input fields to collect the user's age and gender. This data is used to define our target group. During the tasks, our App has three important graphical elements:

- Text label: this contains the text that the user needs to type.
- Edit text: the place where the user types the text.
- Soft keyboard with swift technology enabled/disabled depending on the task.

The app was designed to have no interference of the moderator during the sessions. Additionally, users were asked not to use word suggestion while typing.

EXPERIMENT EXECUTION

We conducted the experiment on two days, we randomly selected twelve students from Bauhaus University, Weimar. After contacting the participants and acquiring their agreement on taking part in the user study, we briefly explained about how the experiment would be conducted and how the tasks would look like. We also explained to them how swipe typing works and let them try it. We had a button to reset the current task in case a problem occurred, avoiding the need to redo the entire experiment from the beginning. When all the tasks were finished, we sent an email with all the data collected (Fig. 2). The entire experiment steps are illustrated in Figure 3.

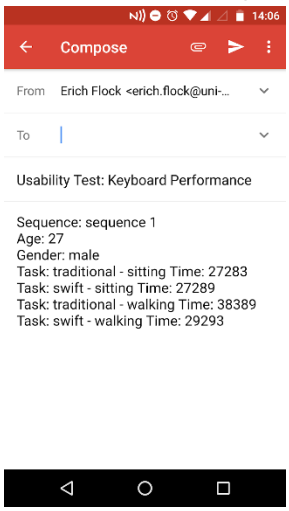

Fig. 2. Report in the end of a user session

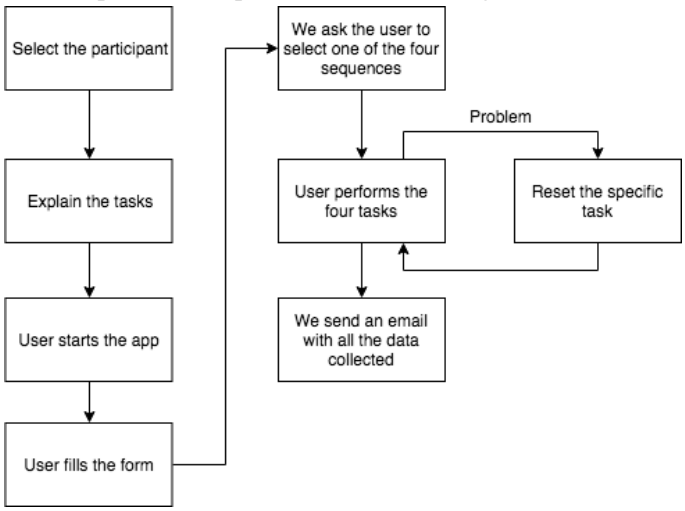

Fig. 3. Test flow

DATA EVALUATION

After gathering all the data from the participants, we analyzed the data and performed significance tests to evaluate it.

- Participants

In total 12 participants (7 male, 5 female) took part in our study. The age distribution is from 20 to 34. Most of them are international students at the Bauhaus University and they use 26-key keyboard on their mobile phones in their daily life.

- Significance test

We mainly implemented the independent T test for significance test. Because of the requirement of the parametric test, we used the Shapiro-Wilk test to test whether the data sets follow the

normal distribution. When the data is distributed normally, we used the independent T test; otherwise, we use Mann-Whitney-U test which is non-parametric significance test.
In our study, we had the following hypotheses which needed testing:

a. For swift-keyboard:

a) H0: the efficiencies of swift-keyboard input during walking and sitting have no significant difference;
H1: the efficiencies of swift-keyboard input during walking and sitting have significant difference;

Table 1: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Swift.Walking	.150	12	.200*	.984	12	.995
Swift.Sitting	.148	12	.200*	.918	12	.271

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 1, the significant values are larger than 0.05, which indicates the time needed for Swift-keyboard during walking and sitting follow the normal distribution.

Table 2: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	1.257	.274	.330	22	.745	6012.000	18229.539	-31793.751	43817.751
	Equal variances not assumed			.330	18.820	.745	6012.000	18229.539	-32167.602	44191.602

From Table 2, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses. In other word, the efficiency for swift-keyboard for walking and sitting have no significant difference.

b. For traditional keyboard:

a) H0: the efficiencies of traditional keyboard input during walking and sitting have no significant difference;
H1: the efficiencies of traditional keyboard input during walking and sitting have significant difference;

Table 3: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Tra.Walking	.201	12	.195	.926	12	.340
Tra.Sitting	.188	12	.200*	.921	12	.293

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 3, the significant values are larger than 0.05, which indicates the time needed for Tradition-keyboard during walking and sitting follow the normal distribution.

Table 4: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	.589	.451	1.251	22	.224	9211.250	7365.113	-6063.059	24485.559
	Equal variances not assumed			1.251	21.948	.224	9211.250	7365.113	-6065.149	24487.649

From Table 4, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiency for traditional-keyboard for walking and sitting have no significant difference.

c. For swift-keyboard and traditional keyboard input during walking:

a) H0: the efficiencies of swift-keyboard input and traditional keyboard input during walking have no significant difference;
H1: the efficiencies of swift-keyboard input and traditional keyboard input during walking have significant difference;

Table 5: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Tra.Walking	.201	12	.195	.926	12	.340
Swift.Walking	.150	12	.200*	.984	12	.995

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 5, the significant values are larger than 0.05, which indicates the time needed for Tradition-keyboard and Swift-keyboard during walking follow the normal distribution.

Table 6:Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	2.027	.169	-1.876	22	.074	-21078.917	11238.107	-44385.325	2227.492

	Equal variances not assumed			-1.876	16.896	.078	-21078.917	11238.107	-44800.385	2642.552

From Table 6, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for traditional-keyboard and swift-keyboard during walking have no significant difference.

d. For swift-keyboard and traditional keyboard input during sitting:

a) H0: the efficiencies of swift-keyboard input and traditional keyboard input during sitting have no significant difference;
H1: the efficiencies of swift-keyboard input and traditional keyboard input during sitting have significant difference;

Table 7: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Tra.Sitting	.188	12	.200*	.921	12	.293
Swift.Sitting	.148	12	.200*	.918	12	.271

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 7, the significant values are larger than 0.05, which indicates the time needed for Tradition-keyboard and Swift-keyboard during sitting follow the normal distribution.

Table 8: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	5.978	.023	-1.505	22	.147	-24278.167	16132.760	-57735.462	9179.129
	Equal variances not assumed			-1.505	13.392	.156	-24278.167	16132.760	-59027.360	10471.026

From Table 8, the significant value is smaller than 0.05, which indicates that we can reject the null hypotheses, so the efficiencies for traditional-keyboard and swift-keyboard during sitting have significant difference.

The average time of traditional-keyboard during sitting is 51244.3ms. As for the swift-keyboard, the average time is 75522.5ms. In conclusion, when users are under stable condition, the input efficiency of traditional-keyboard is higher than that of swift-keyboard.

e. For genders:

a) H0: the efficiencies of swift-keyboard input during walking for male and female have no significant difference;
H1: the efficiencies of swift-keyboard input during walking for male and female have significant difference;

Table 9: Tests of Normality

	Gender	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Swift.Walking	female	.336	5	.067	.881	5	.312
	male	.162	7	.200*	.979	7	.954

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 9, the significant values are larger than 0.05, which indicates the time needed for swift-keyboard during walking for female and male follow the normal distribution.

Table 10: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	.357	.564	.973	10	.353	19577.657	20113.090	-25237.099	64392.414
	Equal variances not assumed			1.004	9.602	.340	19577.657	19507.655	-24133.684	63288.998

From Table 10, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for swift-keyboard during walking for female and male have no significant difference.

b) H0: the efficiencies of swift-keyboard input during sitting for male and female have no significant difference;
H1: the efficiencies of swift-keyboard input during sitting for male and female have significant difference;

Table 11: Tests of Normality

	Gender	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Swift.Sitting	female	.213	5	.200*	.977	5	.920
	male	.234	7	.200*	.877	7	.215

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 11, the significant values are larger than 0.05, which indicates the time needed for swift-keyboard during sitting for female and male follow the normal distribution.

Table 12: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	3.967	.074	.102	10	.921	3305.314	32557.801	-69237.986	75848.615
	Equal variances not assumed			.117	8.103	.910	3305.314	28369.309	-61970.011	68580.640

From Table 12, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for swift-keyboard during sitting for female and male have no significant difference.

c) H0: the efficiencies of tradition-keyboard input during walking for male and female have no significant difference;
H1: the efficiencies of tradition-keyboard input during walking for male and female have significant difference;

Table 13: Tests of Normality

	Gender	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Tra.Walking	female	.230	5	.200*	.948	5	.720
	male	.185	7	.200*	.936	7	.599

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 13, the significant values are larger than 0.05, which indicates the time needed for tradition-keyboard during walking for female and male follow the normal distribution.

Table 14: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	1.928	.195	-1.323	10	.215	-13847.629	10465.889	-37167.082	9471.825
	Equal variances not assumed			-1.399	9.969	.192	-13847.629	9898.964	-35913.068	8217.811

From Table 14, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for tradition-keyboard during walking for female and male have no significant difference.

d) H0: the efficiencies of tradition-keyboard input during sitting for male and female have no significant difference;
H1: the efficiencies of tradition-keyboard input during sitting for male and female have significant difference;

Table 15: Tests of Normality

	Gender	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Tra.Sitting	female	.251	5	.200*	.838	5	.159
	male	.166	7	.200*	.942	7	.653

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 15, the significant values are larger than 0.05, which indicates the time needed for tradition-keyboard during sitting for female and male follow the normal distribution.

Table 16: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	3.892	.077	-1.023	10	.331	-10515.657	10282.434	-33426.347	12395.033
	Equal variances not assumed			-1.187	7.560	.271	-10515.657	8855.492	-31145.388	10114.073

From Table 16, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for tradition-keyboard during sitting for female and male have no significant difference.

f. For the age group:

a) H0: the efficiencies of swift-keyboard input during walking for different age groups have no significant difference;
H1: the efficiencies of swift-keyboard input during walking for different age groups have significant difference;

Table 17: Tests of Normality

	Age_group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Swift.Walking	<=25	.232	6	.200*	.871	6	.231

	>25	.144	6	.200*	.969	6	.884

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 17, the significant values are larger than 0.05, which indicates the time needed for swift-keyboard during walking for different age groups follow the normal distribution.

Table 18: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	.180	.681	-1.904	10	.086	-33850.667	17775.722	-73457.442	5756.109
	Equal variances not assumed			-1.904	9.791	.087	-33850.667	17775.722	-73572.437	5871.104

From Table 18, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for swift-keyboard during walking for different age group have no significant difference.

b) H0: the efficiencies of swift-keyboard input during sitting for different age groups have no significant difference;
H1: the efficiencies of swift-keyboard input during sitting for different age groups have significant difference;

Table 19: Tests of Normality

	Age_group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Swift.Sitting	<=25	.322	6	.052	.840	6	.130
	>25	.174	6	.200*	.965	6	.857

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 19, the significant values are larger than 0.05, which indicates the time needed for swift-keyboard during sitting for different age groups follow the normal distribution.

Table 20: Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper

time	Equal variances assumed	.708	.420	-.794	10	.446	-24737.000	31151.820	-94147.581	44673.581
	Equal variances not assumed			-.794	8.432	.449	-24737.000	31151.820	-95937.464	46463.464

From Table 20, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for swift-keyboard during sitting for different age group have no significant difference.

c) H0: the efficiencies of tradition-keyboard input during walking for different age groups have no significant difference;
H1: the efficiencies of tradition-keyboard input during walking for different age groups have significant difference;

Table 21: Tests of Normality

	Age_group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Tra.Walking	<=25	.338	6	.031	.793	6	.051
	>25	.267	6	.200*	.871	6	.229

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 21, the significant value is smaller than 0.05, which indicates the time needed for swift-keyboard during sitting for different age groups doesn`t follow the normal distribution; so we cannot use T test.

d) H0: the efficiencies of tradition-keyboard input during sitting for different age groups have no significant difference;
H1: the efficiencies of tradition-keyboard input during sitting for different age groups have significant difference;

Table 22: Tests of Normality

	Age_group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Tra.Sitting	<=25	.206	6	.200*	.938	6	.646
	>25	.253	6	.200*	.816	6	.081

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

From Table 22, the significant values are larger than 0.05, which indicates the time needed for tradition-keyboard during sitting for different age groups follow the normal distribution.

Table 23: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
time	Equal variances assumed	1.510	.247	-1.372	10	.200	-13411.667	9775.224	-35192.223	8368.890
	Equal variances not assumed			-1.372	9.126	.203	-13411.667	9775.224	-35478.389	8655.056

From Table 23, the significant value is greater than 0.05, which indicates that we cannot reject the null hypotheses, so the efficiencies for tradition-keyboard during sitting for different age group have no significant difference.

LIMITATIONS

Most people are much more accustomed to typing using the traditional method than swipe typing; therefore, experience is a factor which significantly affects the results of the experiment. However, we could not find any reliable way to control this variable. Besides that, the sample size is not big enough to verify our hypotheses.

DISCUSSION AND CONCLUSION

About the efficiencies of the tradition-keyboard and swift-keyboard, there are no significant difference between in different input conditions, genders and age group. Only when users type the texts during sitting, using traditional keyboard is more efficient than using the swift-keyboard. However, because of the limit of sample size and control of participants` type style, the validity of the experiment and conclusion is not high enough.

FUTURE WORK

Given that users' experience was not fully controlled in the experiment, a different study design that can eliminate the influence of this factor is needed for future experiments.

It is possible that different screen sizes can affect the performance of the typing methods. A study incorporating mobile devices of different screen sizes (mobile phones, tablets and phablets) will help us gain more insight to this factor.

REFERENCES

1. Anson, D., Brandon, C., Davis, A., Hill, M., Michalik, B. and Sennett, C., 2012. Swype vrs. conventional on-screen keyboards: Efficacy compared. In *RESNA Annual Conference*.
2. Cuaresma, J., & MacKenzie, I. S. (2013). A study of variations of Qwerty soft keyboards for mobile phones. *Proceedings of the International Conference on Multimedia and Human-Computer Interaction - MHCI 2013*, pp. 126.1-126.8. Ottawa, Canada: International ASET, Inc..
3. Global mobile OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2017 https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/