University of
St Andrews

Practical 2

CS5040 HCI Principles and Methods

Matriculation ID: 230015450

Question1:

- Issue1: Accessibility

This method of using gestures to control how to select the corresponding letters seems natural, but it does not take into account that some disabled people may not be able to perform this operation with their hands. This needs to be taken into account that different types of disabled people should be matched with corresponding functions.

- Issue2: Accidental touch

Users inadvertently trigger a certain function of the screen or device, or select the wrong corresponding word, instead of what they intended. This may lead to a degradation of user experience and reduce the accuracy and efficiency of interactions.

- Issue3: User adaptation

How to ensure that users can easily learn the operation of new products and enable them to adapt quickly. This is an important factor affecting their experience.

- Issue4: Different application scenarios

Users may use the watch in different environments, so the design needs to consider usability outdoors, in noisy environments, or in low light conditions to ensure it works well in a variety of situations.

Experimental Design

- Setup：In this experiment, I would like to evaluate the usability of the Crownboard watch keyboard in different environments, including indoors, outdoors， to examine its performance in diverse usage scenarios.
- Dependent Variable: Input accuracy and speed
- Independent Variable: Environments
- Participants: Twenty participants voluntarily participated in the study and were recruited through contact information obtained through social media or paid message boards. While maintaining gender balance, each subject will receive a $10 reward. No one reported that their medical condition limited their fine motor skills
- Apparatus: Manual Crownboard
- Material: A simple questionnaire in the beginning that categorize their culture, age, and gender. Consent forms and waivers
- Procedure: Participants were asked to use the Crownboard watch keyboard to perform text input tasks in different environments. Each person was given ten random phrases to input, and the researchers measured their input accuracy and speed.
- Methods: Within-subjects design

- Threats: There may be learning effects that lead to inaccurate results in the post-experiment phase. Mitigate this threat by introducing appropriate practice and familiarity into the experiment. And experimental results may be affected by other factors not considered, such as the noise level, thereby reducing the internal validity of the experiment.

Question2:

Results	Table1	Table2	Table3
SS_{error}	108.8	72.8	592.8571429
SS_{total}	120.9	130.4	1394.666667
SS_{effect}	12.1	57.6	801.8095238
n Participants	10	10	21
m Groups	2	2	3
df_{error}	8	8	18
df_{effect}	1	1	2
MS_{error}	13.6	9.1	32.93650794
MS_{effect}	12.1	57.6	400.9047619
α Confidence Level	0.05	0.05	0.05
F-ratio	0.889705882	6.32967033	12.17204819
Critical value	5.317655072	5.317655072	3.554557146
Significant?	FALSE	TRUE	TRUE

The Gesture experiment conducted resulted in fewer average errors than baseline (2 vs 10respectively). We assumed these errors were normally distributed. Analysis of variance at asignificance level of $\alpha = 0.05$ showed that this difference was statistically significant ($F = 0.89$, $p > 0.05$).

The Eye Tracking experiment conducted resulted in fewer average errors than baseline (2 vs 10 respectively). We assumed these errors were normally distributed. Analysis of variance at a significance level of $\alpha = 0.05$ showed that this difference was statistically significant ($F = 6.33$, $p > 0.05$).

The Keyboard experiment conducted resulted in fewer average errors than baseline (3 vs 21respectively). We assumed these errors were normally distributed. Analysis of variance at asignificance level of $\alpha = 0.05$ showed that this difference was statistically significant ($F = 12.17$, $p < 0.05$).

Qusetion3:

Report

Introduction

In this question the researcher has designed within-subjects performance test to capture the performance levels of 3 mobile keyboard devices. The independent variable contains 3 levels, which are regular, laser projected and VR keyboard. There are 3 dependent variables which are text entry rate (wpm), text error rate (measured in percentage) and preference (scale of 1-7, where 1 is least interest).

Research Method

One of the common ways to interpret this data is to use one-way repeated measure design (Field, 2018), where the independent variable is referred to as the repeated factor (the types of keyboards) while the dependent variable is referred to as the repeated measure (type speed, typing errors). In this scenario, typing speed, typing error and preference was measured repeatedly over the different types of keyboards and the subjects serves as their own controls (Lomax & Hahs-Vaughn, 2012). This method allows the individual differences to be negated within the variability of the error terms (i.e., categorizing each subject to their own blocks).

Data analysis and results

A reasonable hypothesis here is that the performance and preferences will vary considerably across the 3 keyboards. If the results are significant, pairwise differences will be explored in means utilizing Bonferroni-adjusted pairwise comparisons.

Within-Subjects Factors

Measure: MEASURE_1

keyboard	Dependent Variable
Regular	TextEntryRateReg
Laser	TextEntryRateProj
VR	TextEntryRateSlide

Descriptive Statistics

	Mean	Std. Deviation	N
Regular(wpm)	36.1875	8.71947	16
Laser(wpm)	21.5625	2.55522	16
VR(wpm)	42.3750	8.80057	16

Table.1 Table of mean values

By calculating the mean of the input rates, it can be clearly seen the VR keyboard is the fastest (Mean=42.3750) followed by regular (Mean=36.1875) then the slowest with laser projected (Mean=21.5625). From the multivariate test, we can assume that the test results are independent of multivariate normality. When sphericity plays an important role, univariate tests are considered more significant. However, multivariate testing is more powerful when sphericity is not considered. This report will discuss multivariate tests first and then discuss sphericity later.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
keyboard	Pillai's Trace	.902	64.285[b]	2.000	14.000	<.001
	Wilks' Lambda	.098	64.285[b]	2.000	14.000	<.001
	Hotelling's Trace	9.184	64.285[b]	2.000	14.000	<.001
	Roy's Largest Root	9.184	64.285[b]	2.000	14.000	<.001

a. Design: Intercept
Within Subjects Design: keyboard

b. Exact statistic

Table 2. Multivariate Test (Text Entry)

Using the Wilk's Lambda from SPSS, it is observed that there is a significant difference in the means. Wilk's Lambda = 0.098, $F = 64.285$, $p < 0.0001 < \alpha = 0.05$, therefore we can reject the null hypothesis.

Tests of Within-Subjects Effects

wpm: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
keyboard	Sphericity Assumed	3655.125	2	1827.562	31.386	<.001
	Greenhouse-Geisser	3655.125	1.476	2476.606	31.386	<.001
	Huynh-Feldt	3655.125	1.598	2287.081	31.386	<.001
	Lower-bound	3655.125	1.000	3655.125	31.386	<.001
Error(keyboard)	Sphericity Assumed	1746.875	30	58.229		
	Greenhouse-Geisser	1746.875	22.138	78.909		
	Huynh-Feldt	1746.875	23.972	72.870		
	Lower-bound	1746.875	15.000	116.458		

Table 3. Univariate repeated measures ANOVA (Text Entry)

When assuming sphericity, there will be a potential risk of committing Type I error, in other words, incorrectly rejecting the null hypothesis when there is no effect. To solve this problem, one can adopt the Greenhouse-Geisser method to correct for sphericity violations or abandon the assumption of sphericity and rely solely on multivariate tests. This is a way of handling data when faced with sphericity violations.

In repeated measures experiments, the test of sphericity is used to check whether co-variance between pairs of conditions is equal (Field, 2018). Sphericity is heavily based on the differences of the variance scores between datasets.

Mauchly's Test of Sphericity[a]

wpm: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
keyboard	.645	6.142	2	.046	.738	.799	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: keyboard

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 4. Mauchly's Sphericity Test (Text Entry)

When the epsilon value is equal to 1, it indicates that sphericity is satisfied. More specifically, in our tests the value was low at 0.738, which deviates from sphericity, and this parameter was then calculated to adjust the degrees of freedom in Table 3. The significant value of Mauchly's test is 0.046, which indicates that our sphericity is not satisfied (Pituch & Stevens, 2016). If the Greenhouse–Geisser estimates of sphericity is below 0.75 or when there is uncertainty about sphericity, it is advisable to apply the Greenhouse–Geisser correction (Field, 2018). Our new data from Greenhouse-Geissier shows of $p=0.046<0.05$, which agrees with Wilk's Lambda in rejecting the null hypothesis. Using the same approach, we will conduct statistical significance tests of text error rates as well as preferences and link them to draw conclusions.

In the experiment error rate testing, the data shows as below:

Within-Subjects Factors

Measure: MEASURE_1

errorrate	Dependent Variable
regular	TextErrorRate Reg
laser	TextErrorRate Proj
VR	TextErrorRate Slide

Descriptive Statistics

	Mean	Std. Deviation	N
TextErrorRateReg	1.44	.512	16
TextErrorRateProj	8.69	1.537	16
TextErrorRateSlide	4.56	1.413	16

Table5. Table of mean values

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
errorrate	Pillai's Trace	.951	137.037[b]	2.000	14.000	<.001
	Wilks' Lambda	.049	137.037[b]	2.000	14.000	<.001
	Hotelling's Trace	19.577	137.037[b]	2.000	14.000	<.001
	Roy's Largest Root	19.577	137.037[b]	2.000	14.000	<.001

a. Design: Intercept
Within Subjects Design: errorrate

b. Exact statistic

Table 6. Multivariate Test (error rate)

Tests of Within-Subjects Effects

errorrate: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
errorrate	Sphericity Assumed	423.167	2	211.583	172.330	<.001
	Greenhouse-Geisser	423.167	1.938	218.346	172.330	<.001
	Huynh-Feldt	423.167	2.000	211.583	172.330	<.001
	Lower-bound	423.167	1.000	423.167	172.330	<.001
Error(errorrate)	Sphericity Assumed	36.833	30	1.228		
	Greenhouse-Geisser	36.833	29.071	1.267		
	Huynh-Feldt	36.833	30.000	1.228		
	Lower-bound	36.833	15.000	2.456		

Table 7. Univariate repeated measures ANOVA (error rate)

Mauchly's Test of Sphericity[a]

errorrate: MEASURE_1

					Epsilon[b]		
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
errorrate	.968	.455	2	.797	.969	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: errorrate

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 8. Mauchly's Sphericity Test (error rate)

When the epsilon value is equal to 1, P>0.05, it indicates that sphericity is satisfied. The next step is to look at the results of the hypothesis sphericity in the test within-subject effects test. The data shows that F=172.33, P<0.001, it Indicates significant differences between groups and rejects null hypothesis.

In the experiment preference level testing, the data shows as below:

Within-Subjects Factors

Measure: MEASURE_1

perfer	Dependent Variable
regular	PreferredUsingReg
laser	PreferredUsingProj
VR	PreferredUsingSlide

Descriptive Statistics

	Mean	Std. Deviation	N
PreferredUsingReg	3.25	.931	16
PreferredUsingProj	1.56	.629	16
PreferredUsingSlide	4.13	1.204	16

Table9. Table of mean values

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.
perfer	Pillai's Trace	.784	25.458[b]	2.000	14.000	<.001
	Wilks' Lambda	.216	25.458[b]	2.000	14.000	<.001
	Hotelling's Trace	3.637	25.458[b]	2.000	14.000	<.001
	Roy's Largest Root	3.637	25.458[b]	2.000	14.000	<.001

a. Design: Intercept
Within Subjects Design: perfer

b. Exact statistic

Table 10. Multivariate Test (preference)

Tests of Within-Subjects Effects

perfer: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
perfer	Sphericity Assumed	54.292	2	27.146	27.412	<.001
	Greenhouse-Geisser	54.292	2.000	27.148	27.412	<.001
	Huynh-Feldt	54.292	2.000	27.146	27.412	<.001
	Lower-bound	54.292	1.000	54.292	27.412	<.001
Error(perfer)	Sphericity Assumed	29.708	30	.990		
	Greenhouse-Geisser	29.708	29.997	.990		
	Huynh-Feldt	29.708	30.000	.990		
	Lower-bound	29.708	15.000	1.981		

Table11. Univariate repeated measures ANOVA (preference)

Mauchly's Test of Sphericity[a]

perfer: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
perfer	1.000	.001	2	.999	1.000	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: perfer

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Table 8. Mauchly’s Sphericity Test (preference)

When the epsilon value is equal to 1, P>0.05, it indicates that sphericity is satisfied. The next step is to look at the results of the hypothesis sphericity in the test within-subject effects test. The data shows that F=27.412, P<0.001, it Indicates significant differences between groups and rejects null hypothesis.

Limitation

One limitation of the single factor repeated measures analysis in this study is the assumption of normality. Although it is common practice to assume normality in statistical analyses, the robustness of repeated measures ANOVA relies on this assumption. If the data deviate significantly from a normal distribution, it may affect the accuracy and reliability of the results. And the number of samples in this experiment is too small. If the analysis of assumes are complete normality, it may cause bias. Therefore, it is necessary to calculate the p value based on the normal distribution test, test the significance level, and analyze whether the assumption of normal distribution is met.

Reference

Field, A. P. (2018). *Discovering statistics using Ibm Spss statistics.* SAGE Publications.

Lomax, R. G., & Hahs-Vaughn, D. L. (2012). *An introduction to statistical concepts.* Routledge.

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the Social Sciences: Analyses with Sas and Ibm's Spss.* Routledge.