

ANALYSIS OF USABILITY OPCS SOFTWARE: A TEST TO THE STUDENT

Rizky Junianto¹, Laurence¹, Agustina Christiani¹, Helena J. Kristina¹

¹Department of Industrial Engineering, University of Pelita Harapan, Lippo Village Tangerang, Indonesia
laurence.uph@gmail.com

ABSTRACT

The research is to know the usability of new software implemented in the university. Previously, the system used in-house developed software called SLIM (Student Information Service System) before turn to Enterprise Resource Planning Oracle PeopleSoft Campus Solution (OPCS). Students are selected to participate in the survey and have to complete various tasks. Several variables of usability are observed such as learnability, efficiency, memorability, error, and satisfaction. According to the result of the experiment from 51 respondents, it is shown that that the usability level of the software is 0.70 (Good). The value of each dimensions are followed: learnability 0.60 (Good), efficiency 0.59 (moderate), errors 0.85 (Excellent), memorability 0.82 (Excellent), and for satisfaction is 0.71(Good).

Keywords : Usability, learnability, efficiency, memorability, error, satisfaction

1. INTRODUCTION

In integrated system, every components work closely to achieve the organization's objective. The role of computer usage as part of the system in digital era is not debatable that its influence is not only to personal life of human but also to the company of the way they run daily routine. Human interaction with computer becomes more intense, as the developer build the user friendly software that enable user to complete the task. However the approach seems not applied to all software, so that the interesting point is whether user should adapts with the software or tailor made one is better.

2. METHODOLOGY

The target of research can be divided depends on the user level in the university. For instance, the object of the survey is student, lecturer, or academic adviser. With the access menus that vary between the users, the type of the questions will be different among them. Although the frequent usage of the software is not daily usage, however the number of students is significant and they are one of the stakeholders that will feel the impact of changes. It is decided in the initial research that the student will be surveyed to know

their opinion about the software. The survey will be conducted into two stages, which measure error and efficiency for the first part, learnability, satisfaction and memorability for the last part. The quantitative value is then associated with qualitative value such as good, bad, excellent to describe the usability of the software. Table 1 exhibits the five dimension and its criterion used in the experiments.

Table 1. Five Dimensions and Measurement

Dimension	Object of Measurement
Error	Number of mistakes for each question.
Efficiency	Time and number of click compared to standard.
Learnability	Time difference between first and last experiment
Memorability	Slack of number of click between two tests
Satisfaction	Result from survey

3. RESULTS

There is some important information that the student need to know, for example the cumulative grade or the subjects already taken. In the first survey of 51 students, they were given five questions about the activity that the students usually do in the beginning or at the end of academic term. The academic activity that the every student

needs to do is how to enroll, drop, swap the subjects, checks score of subject and grade point average.

3.1. Error

Error dimension is defined as non-performed or unfinished result or fail to be solved which is calculated by counting the number of mistakes of the answer given. The data is shown in table 1

Table 2. Answer of Five Questions

	Question				
Answer	Q 1	Q2	Q3	Q4	Q5
Wrong	2	5	10	6	14
Right	49	46	41	45	37

3.2. Efficiency

Efficiency unit is taken by calculating time and number of clicks by respondent to answer the five questions. The result will be compared with the standard, which is defined as the time and number of clicks of the task’s creator. Table 3 illustrates how to record the result of the survey. For instance, time for respondent 1 is recorded to 0 and number of clicks is 1. It means he finished the task faster than time standard, but for number of click, he did one more step. For respondent 2, he completed the question in 46 seconds in the same number of clicks as the standard.

Table 3. Comparison to Standard

Respondent	Difference	
	Time	Click
1	0	1
2	46	0

The mean, standard time and clicks is described in table 4. The data is gathered in two different classes, where the result not much different in general.

Table 4. Standard and Average Data of First Experiment

Question	Standard		Class 1		Class 2	
	Time	Click	Time	Click	Time	Click
1	20	5	64.14	2.06	76.86	2.43
2	103	14	84.16	5.44	159	3.29
3	75	12	67.41	4.9	65	3.83
4	19	4	30.81	2.68	26.62	0.92
5	24	4	7.43	12.52	15.92	12.67

Second experiment runs two weeks after the first one with the same questions and same

type of data. The different is that the respondents are not allowed to see the handout. Table 5 illustrates the result the mean of data of second experiment.

Table 5 Standard and Average Data of Second Experiment

Question	Mean	
	Time	Click
1	44,69	6,41
2	126,25	13,72
3	84,47	13,50
4	43,50	4,16
5	29,09	3,63

Only 14 respondents of the total sample can consistently provide the right answer from both experiments. This group of people is the same persons. Figure 1 exhibits the comparison of total number of right answer for 5 questions in the first and second experiments. Number of click and time recorded will be used to find value of learnability and memorability consecutively. The slope of the curve is decreasing, however it does not mean that the question intentionally given from the easiest to the most difficult level.

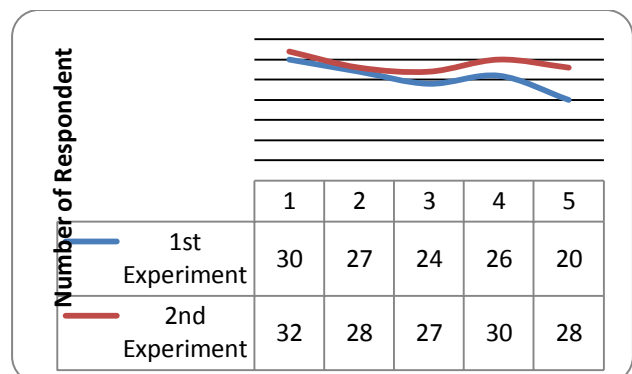


Figure.1. Comparison of the first and second experiment

If the both of experiments are compared, it is known that there is increase number of successive students that answer correctly.

3.3. Learnability

Learnability is one important factor that needs to be measured to know whether the software is easy to learn or not. The value can be calculated by finding the time difference between both experiments.

Table 6. Comparison of Mean and Standard Deviation From First and Second Experiment.

Question	Mean (seconds)		Standard Deviation (seconds)	
	Experiment		Experiment	
	1	2	1	2
1	52.07	50.86	20.50	33,78
2	170.00	142.14	49.38	46.47
3	141.00	85.36	44.13	32.93
4	31.43	46.50	39.54	38.21
5	26.71	34.14	13.53	21.54

Table 6 shows that the mean time to finish the question 1 to 3 in second trial is less than the first one. This is contrast with the question 4 to 5. Even though the difference less than one minute, but it is necessary to know whether they differ one another. Hypothesis test to know whether there is significant different of mean time between both experiments.

H_0 : Mean of both experiments is not significantly different, $\mu_1 = \mu_2$

H_1 : There is significant different between first and the second experiment, $\mu_1 \neq \mu_2$

If the significant value is less than 0.05, then it can be concluded that the experiment differ one another. Table 7 shows the result of mean test.

Table 7. Mean Test

Question	Significant Value	Conclusion
1	0.315	Accept H_0
2	0.124	Accept H_0
3	0.001	Accept H_1
4	0.177	Accept H_0
5	0.001	Accept H_1

It can be concluded that the question no 3 and 5 significantly different because the value is less than 0.05. Question number 3 is about how to swap the subject without using drop subject menu. This menu is not familiar with the student, because they never see this option in the previous software. Question number 5 asks the student the way to see grade point average. In this case, the students are more familiar and remember when they try in the second runs. The Wilcoxon test in figure 3 shows that both

experiments are not the same, with the result significant value 0.002 and the decision is to reject null hypothesis.

Null Hypothesis	Test	Sig.	Decision
1 The median of differences between T_L_1 and T_L_2 equals 0.	Related-Samples Wilcoxon Signed Rank Test	,002	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Figure 3. Wilcoxon Test for Difference

3.4. Memorability

Memorability refers to how the user shows no difficulties to use the software after some of breaking period. Numbers of clicks by users that have been recorded of first and second experiment are compared. Table 8 shows the mean of both experiments for case 1 until 5.

Table 8 Number of Clicks

Question	Experiment 1		Experiment 2		Slack
	Mean	Stdev	Mean	Stdev	
1	5.86	1.88	6.14	0.36	0.28
2	18.29	3.58	15.14	1.41	-3.15
3	17.57	3.50	14.71	0.99	-2.86
4	5.14	2.77	4.57	1.50	-0.57
5	4.36	0.74	4.29	0.61	-0.07

The slack value varies from -3.15 to 0.28. Negative numbers exist four times which show that there is reduction value of the last experiment compare to the previous one. Compared to the first experiment, the standard deviation for experiment two show the steady number. The hypothesis testing is run to know whether there is significant different between number of click from both samples.

H_0 : There is no significant difference between number of click from first and second experiment, $\mu_1 = \mu_2$

H_1 : There is significant difference between number of click from first and second experiment $\mu_1 \neq \mu_2$

Data is different if the sig value is less than 0.05. Accept null hypothesis, if the sig value is greater than 0.05.

Table 9. Significant Test

Q	Sig Value	Decision	Conclusion
1	0.19	Accept Ho	No significant different
2	0.018	Accept H1	There is significant different
3	0.016	Accept H1	There is significant different
4	0.588	Accept Ho	No significant different
5	0.679	Accept Ho	No significant different

Table 7 shows that no significant different of click's number for question 1, 4, and 5, while the contrast result show for question 2 and 3. The total significant value after calculated using the software is 0.007 which is less than 0.05. Based on this result, it can be concluded that both the test are not diverse in number of clicks.

3.5. Satisfaction

After they finish the task, the respondents need to fill the survey consist of 12 questions from scale 1 strongly disagree to 5 strongly agree about their opinion when using this software. The students usually need information about their grade of subject, final GPA, how to enroll the class and to check the schedule. Basically the questions ask them about their opinion of menu navigation how easy to operate.

Table 10. Survey Result

S	Question											
1	1	2	2	1	3	2	2	3	2	2	3	4
2	12	5	7	6	5	8	8	9	13	12	8	9
3	12	14	20	20	10	11	9	16	11	12	15	11
4	7	10	3	5	13	9	10	4	6	6	5	8
5	0	1	0	0	1	2	3	0	0	0	1	0

Majority of the respondent choose 3 which means neutral position. It is understandable because some of them feel difficult to become accustomed with new routing. However the final calculation for satisfaction is 0.71

3.6. Usability Level

Here normalization is used to get the single value of all the aspect, in order the value of usability software can be judge good or not. Efficiency has two indicators, which are time and number of click. Normalization value for time aspect is 0.50, and for number of click is 0.68. The average value of two indicators, which result 0.59 will be used to represent

efficiency dimension. Figure 2 display the spider graph of five dimensions.

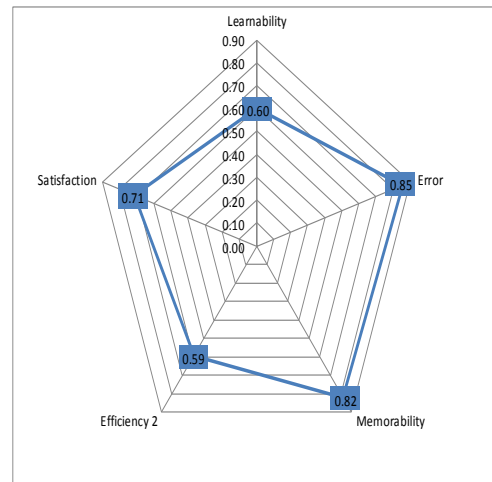


Figure 2. Value of Each Dimension

This value is located in the range moderate which between 0.4 until 0.6. Value after normalization of error dimension is 0.85. This number is lies in the category good. Memorability in these tests is the comparison of number of clicks for first and second test. The mean value is 0.82, which belong to category very good. Next is learnability which the result is from time comparison from first and second experiment. The outcome is 0.60 which can be considered good. Satisfaction rate from the survey is 0.71. This level is considered good rate, which is the same as learnability dimension. To know the overall value, it is required find the average from all aspects. The average is 0.70 which can be grouped as good.

4. DISCUSSION

Category moderate in time show that the respondents still confuse how to solve the task given. It is normal because previous software that is used by students is straightforward compared to this one. For instance, most of them need more time to adapt with new enrollment menu. When the time required increase, it is directly related with number of click. In the end, both parameters will be above the standard given.

In spite of growth in time occur, the number of error is not substantial. Figure 3 shows that the percentage people who

made maximum 1 mistake is cover about 80% of the total sample. It means most of the user can solve the problem, even though it takes longer time. From the satisfaction survey, it can be drawn that regardless of longer time required, the respondents were not against using the software or over negative reaction.

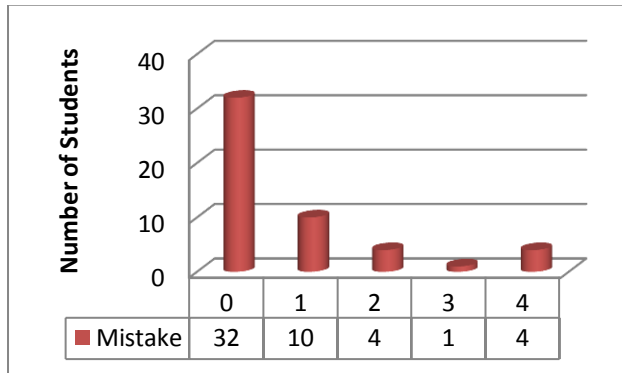


Figure 3. Error Each Person

Human is naturally born with the ability to memorize something. It is well known that memory ability is divided into two category, short term and long term memory. The result in figure 4 exhibits that negative value is dominated by answer number two and three. The data recorded is maximum 25 clicks by respondent 1 in the initial test and decrease to 14 clicks two weeks later. Same amount reduction applied to 14th respondent while answering question number 3. It is enough with the negative number with amount of 40% to draw conclusion that the step to complete the task is able to be remembered.

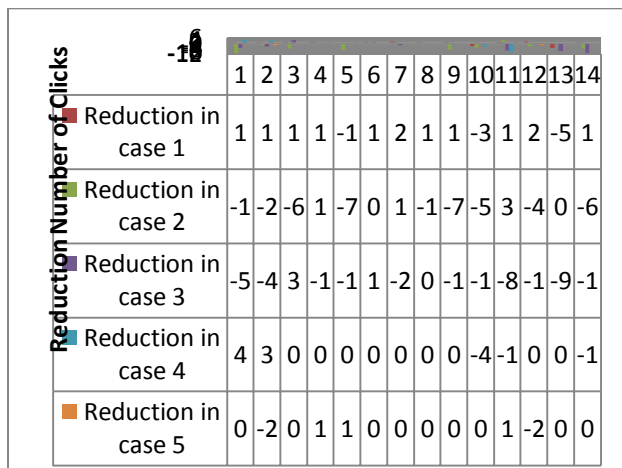


Figure 4. Reduction Number of Clicks

5. CONCLUSION

Efficiency is measured by the time and number of clicks in the first experiment compared with the standard. Result is 0.59 which can be categorized as moderate. Error usability is 0.85 which is considered excellent. The value is result based on answer of the students. Value of dimension of learnability, memorability, and satisfaction as followed: 0.6 (good), 0.82(excellent), and 91.5. With that level of satisfaction, it means the user feel satisfy with the software, however when compared with overall usability value 0.70 after normalization, the category is good.

6. REFERENCES

- (a) Chapanis, Alphonse. (1993). *Applied Experimental Psychology: Human Factors in Engineering Design*. United States: John Wiley.
- (b) G. Perlman, G.K. Green, M.S. Wogalter. *Human Factor Perspectives on Human-Computer Interaction: Selection from proceeding of human factors and ergonomics society annual meeting 1983-1994*. Santamonica: CA: Human Factors & Ergonomics Soc.
- (c) ISO. (1997). "ISO 9241-11: Draft International Standard on Ergonomics Requirement for Office Work with Visual Display Terminals (VDT)." Dalam *Part 11: Guidance of Usability*.
- (d) Sastramihardja, Husni S. (1999). Perancangan Kerja dalam Perangkat Lunak Interaktif." *Jurnal Teknik Informatika ITB* Volume 1 Nomor 1.
- (e) J.Rubin, Dana Chisnell. (2008). *"Handbook of Usability Testing"*. Indianapolis, Indiana: Wiley Publishing, Inc..
- (f) Santoso, Singgih. (2007). *Menguasi Statistik di Era Informasi dengan SPSS 15*. Jakarta: PT. Elex Media Komputindo,.
- (g) Ms and Mc' Cornick, Sanders. (1993). *Human Factors in Engineering and Design*. New York: Mc. Graw-Hill Book Co.
- (h) Nielsen, Jacob. *Usability 101: Introduction to Usability*.

- <http://www.useit.com/alertbox/20030825.html>
- (i) OPCS Universitas Pelita Harapan.
<http://web.academic.uph.edu/>
 - (j) Salim, Siti Satwa, dan Thian Kian Chiew. (2003). "WebUse: Website Usability Evaluation Tool." *Malaysian Journal of Computer Science*.
 - (k) White, Catherine M. (1999). *Ergonomics: What is it? Clearing Away the Confusion*. Michigan Gamma, 1999.
 - (l) Yusdhistira. *Usability*. (2007).
[http://mti.ugm.ac.id/~yudhistira/Resource MTI/UAS/IMK\(22-07-07\)/11%20-%20Usability.pdf](http://mti.ugm.ac.id/~yudhistira/Resource%20MTI/UAS/IMK(22-07-07)/11%20-%20Usability.pdf).
 - (m) Sevilla, Consuelo G. et. al (2007). *Research Methods*. Rex Printing Company. Quezon City
 - (n) Montgomery, Douglas C. (2009). *Statistical Quality Control: A Modern Introduction*, 6th ed. New York: John Wiley & Sons.
 - (o) Nielsen, Jakob. (1993). *Usability Engineering*. United Kingdom: Academic Press Limited.
 - (p) Peterson, Robert A. (1994). A Meta-Analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*. *Journal of Consumer Research*, Vol. 21, No. 2 (Sep., 1994)